Instability of nondiscrete free subgroups in Lie groups

Alexey Glutsyuk

March 21, 2008

Abstract

Consider a nondiscrete free subgroup with two generators in a Lie group.

We study the following question stated by Étienne Ghys: is it always possible to make arbitrarily small perturbation of the generators of the free subgroup in such a way that the new group formed by the perturbed generators be not free? In other terms, is it possible to approximate the pair of generators of free subgroup by pairs with relations?

We prove the positive answer.

We study the question (also stated by Ghys) on the best approximation rate in terms of the minimal length of relation in the approximating group. We give an upper bound of the best approximation rate as an exponent of minus the κ- th power of the minimal relation length, $0.19 < \kappa < 0.2$, see (1.6).

We give a survey of related results and open questions.

Contents

1 Introduction and the plan of the paper **2**

1.1 Main result: instability of liberty. Plan of the paper 2
1.2 Approximations by values of words. 4
1.3 Approximations by groups with relations 9
1.4 Historical remarks and further open questions 10
1.5 A simple proof of Theorem 1.1 for $G = PSL_2(\mathbb{R})$ 11
1.6 Case of group $Aff_+(\mathbb{R})$. 11
1.7 Generalization in the case of semisimple Lie group with irreducible adjoint . . 12
1.8 Background material 1: Lie groups, basic definitions and properties. 12
1.9 Background material 2: semisimple Lie algebras and root decomposition . . . 15
1.10 Background material 3: proximal elements 17

2 Proof of Theorems 1.1 and 1.33 for semisimple Lie groups with irreducible adjoint and proximal elements **19**

2.1 Motivation and the plan of the proof 19
2.2 Dynamics of commutator. Proof of Proposition 2.2 23
2.3 The diffeomorphicity of Ψ. Proof of Lemma 2.6 modulo the Main Technical Lemma 25
2.4 Independent eigenvalues. Proof of Main Technical Lemma 2.5 26

3 Case of semisimple Lie groups with irreducible adjoint and without proximal elements 30
3.1 The plan of the proof of Theorem 1.33 30
3.2 Complex eigenvalues of adjoint and $\mathbb{C}$-1-proximal elements 36
3.3 Dynamics of commutator. Proof of Proposition 3.8 38
3.4 Nondegeneracy of the derivative. Proof of Lemma 3.11 39

4 Proof of Theorem 1.1 for arbitrary Lie group 48

5 A short proof of Theorem 1.1 for dense subgroups in $G = PSL_2(\mathbb{C})$ 51

6 Approximations by pairs with relations. Proof of Theorem 1.29 53
6.1 Case of semisimple Lie group G with irreducible adjoint 53
6.2 Case of arbitrary semisimple Lie group 55
6.3 Case of non-semisimple Lie group 59

7 Approximability of semisimple Lie groups 60
7.1 The weak Solovay-Kitaev inequality. Proof of Lemma 1.25 60
7.2 Approximations with weak Solovay-Kitaev property. Proof of Theorem 1.26 . 62

8 Acknowledgments 69

1 Introduction and the plan of the paper

1.1 Main result: instability of liberty. Plan of the paper

Let G be a nonsolvable Lie group. It is well-known (see [11]) that almost each (in the sense of the Haar measure) pair of elements $(A, B) \in G \times G$ generates a free subgroup in G. At the same time in the case, when G is connected and semisimple, there is a neighborhood $U \subset G \times G$ of unity in $G \times G$ where a topologically-generic pair $(A, B) \in U$ generates a dense subgroup: the latter pairs form an open dense subset in U. This was proved in [4].

The pairs generating groups with relations form a countable union of surfaces (relation surfaces) in $G \times G$. We show that the relation surfaces are dense in U.

The main result of the paper is the following

Theorem 1.1 *Any nondiscrete free subgroup with two generators in a nonsolvable Lie group G is unstable. More precisely, consider two elements $A, B \in G$ generating a free subgroup $\Gamma =< A, B >$. Let Γ be not discrete. Then there exists a sequence $(A_k, B_k) \to (A, B)$ of pairs converging to (A, B) such that the corresponding groups $< A_k, B_k >$ have relations: there exists a sequence $w_k = w_k(a, b)$ of nontrivial abstract words in symbols a, b (and their inverses a^{-1}, b^{-1})* [1] *such that $w_k(A_k, B_k) = 1$ for all k.*

Remark 1.2 The condition that the subgroup under consideration be nondiscrete is natural: one can provide examples of discrete free subgroups of $PSL_2(\mathbb{C})$ (e.g., the Schottky group, see [3]) that are stably free, i.e., remain free under any small perturbation of the generators.

[1] Everywhere in the paper, by a word in given symbols we mean a word in the same symbols and their inverses

Remark 1.3 The closure of a nondiscrete subgroup in a Lie group is a Lie subgroup of positive dimension (see [21], p.42). Therefore, in Theorem 1.1 without loss of generality we assume that the subgroup $< A, B >\subset G$ under consideration is dense in G.

The question of instability of nondiscrete free subgroups was stated by É Ghys, who also suggested to study the best rate of approximation of the pair (A, B) by pairs having a relation of a length no greater than a given l (in analogy with the approximations of irrational number by rationals, where the best approximation rate is well-known; it is achieved by continued fractions. In our situation the pair (A, B) plays the role of an irrational number, the pairs with relations play the role of rationals.) We prove an upper bound of the best approximation rate (Theorem 1.29 and Corollaries 1.30, 1.31 stated in 1.3 and proved in 6 and 1.3).

The proof of Corollary 1.30 uses Theorem 1.16 (stated in 1.2), which deals with a semisimple Lie group and a pair (A, B) of its elements generating a dense subgroup (briefly called an irrational pair). It provides an upper bound for the rate of approximations of the elements of the unit ball in the Lie group by words in (A, B) satisfying a bound of derivatives. These and related results and open problems are discussed in Subsections 1.2-1.4.

Theorem 1.16 follows (see 1.2) from Lemma 1.25 and Theorem 1.26, both stated in 1.2 and proved in 7. Theorem 1.26 proves the statement of Theorem 1.16 for a Lie group whose Lie algebra satisfies the so-called weak Solovay-Kitaev inequality (see Definition 1.23). This inequality means a decomposition (with estimate) of each element of a Lie algebra as a sum of two Lie brackets. Lemma 1.25 shows that the latter inequality holds true for any semisimple Lie algebra.

Theorem 1.21 (recalled in 1.2 and proved by R.Solovay and A.Kitaev in [6, 15, 17]) concerns the Lie groups whose Lie algebras satisfy the (strong) Solovay-Kitaev inequality (see Definition 1.17). This inequality says that each element of a Lie algebra is a Lie bracket (with estimate). For these Lie groups Theorem 1.21 provides an upper bound for the rate of approximations of its elements in the unit ball by words in a given irrational pair of elements. The bound given by Theorem 1.21 is stronger than that in Theorem 1.16. Corollary 1.31 follows (see 1.2) from Theorems 1.21, 1.29 and Remark 1.22 (whose statement is proved at the end of Section 7).

Remark 1.4 In the case, when the Lie group under consideration is $PSL_2(\mathbb{R})$, Theorem 1.1 easily follows from the density of the elliptic elements of finite orders in an open domain of $PSL_2(\mathbb{R})$: the proof is given in Subsection 1.5. The case of $PSL_2(\mathbb{C})$ is already nontrivial (in some sense, this is a first nontrivial case). In this case the previous argument cannot be applied, since the elliptic elements in $PSL_2(\mathbb{C})$ are nowhere dense. At the same time, there is a short proof of Theorem 1.1 for dense subgroups in $PSL_2(\mathbb{C})$ that uses holomorphic motions and quasiconformal mappings. We present it in Section 5.

Theorem 1.1 is proved in Sections 2-4. First we prove it (in Section 2) for semisimple Lie groups with irreducible adjoint and proximal elements - a class of Lie groups containing all the groups $SL_n(\mathbb{R})$ and more generally, all the simple split groups (see [21], p.288). A reader can read the proofs in Section 2 assuming everywhere that $G = SL_n(\mathbb{R})$.

Then in Section 3 we prove it for all the other semisimple Lie groups with irreducible adjoint. Afterwards in Section 4 we deduce the statement of Theorem 1.1 for arbitrary Lie group, using the classical radical and decomposition theorems for Lie algebras (see [21], pp. 60, 61, 151; they are briefly recalled in Subsection 1.8).

In 1.7 we formulate a more general Theorem 1.33 in the case of a semisimple Lie group with irreducible adjoint representation. We deduce Theorem 1.1 from it at the same place. We prove Theorem 1.33 for groups with proximal elements in Section 2 and for groups without proximal elements in Section 3.

The definition of proximal element and basic properties of groups with proximal elements will be recalled in 1.10.

In 1.4 we present a brief historical overview and some open problems.

In 1.6 we give a proof of a simplified analogue (Proposition 1.32) of Theorem 1.1 for the simplest solvable noncommutative Lie group $Aff_+(\mathbb{R})$, which is the group of orientation-preserving affine transformations of the real line. (The author is sure that Proposition 1.32 is well known to the specialists.) The proof gives a simple illustration of the basic ideas used in the proof of Theorem 1.1.

The basic definitions concerning Lie groups (adjoint representation, (semi) simple groups, etc.), which will be used through the paper (mostly in proofs), are recalled in 1.8 and 1.9.

1.2 Approximations by values of words.

Definition 1.5 Let G be a Lie group. We say that a pair $(A, B) \in G \times G$ is *irrational*, if it generates a dense subgroup in G.

Proposition 1.6 *Let G be a semisimple Lie group. The set of irrational pairs in $G \times G$ is open. More generally, the set of M- ples of elements of G generating dense subgroups is open in the product of M copies of G.*

Proof We prove the statement of the proposition for pairs: for M- ples the proof is analogous. Let $(A, B) \in G \times G$ be an irrational pair. We have to show that there exists its neighborhood $V \subset G \times G$ such that each pair $(A', B') \in V$ is irrational. Let $G_0 \subset G$ be the unity component of G. Recall that there exists a neighborhood $U \subset G_0 \times G_0$ of unity where an open and dense set of pairs generate dense subgroups in G_0 (see the beginning of the paper and [4]). Thus, there exists an open subset $U' = U_1 \times U_2 \subset U$ such that each pair in U' generates a dense subgroup in G_0. There exist words w_1 and w_2 such that $w_j(A, B) \in U_j$, $j = 1, 2$. By continuity, there exists a neighborhood V of (A, B) such that for any $(A', B') \in V$ one has $w_j(A', B') \in U_j$, and thus, the subgroup generated by $w_j(A', B')$ is dense in G_0 by definition. The ambient subgroup generated by (A', B') is dense in G, since its closure contains G_0 (the previous statement) and each connected component of G contains an element of $< A', B' >$. (The latter fact holds true for the subgroup $< A, B >$ (which is dense) and remains valid for $< A', B' >$ by continuity.) Thus, each pair $(A', B') \in V$ is irrational. The proposition is proved. □

Let us recall the following well-known

Definition 1.7 Given a metric space E, a subset $K \subset E$ and a $\delta > 0$. We say that a subset in E is a δ- net on K, if the union of the δ- neighborhoods of its elements covers K, and all these neighborhoods do intersect K.

Remark 1.8 A δ- net on K is always contained in the δ- neighborhood of K.

Everywhere below (whenever the contrary is not specified) for any given point a of the space $\mathbb{R}^n$ (or of a Lie group G equipped with a metric) we denote

$D_r(a)$ the ball centered at a of radius r, $D_r = D_r(0) \subset \mathbb{R}^n$ (respectively, $D_r = D_r(1) \subset G$).

Everywhere below whenever we say about a distance on a Lie group, we measure it with respect to a given left-invariant Riemann metric on the group (if the contrary is not specified). We use the following property of left-invariant distance.

Proposition 1.9 *Let $\delta_1, \delta_2 > 0$, G be a Lie group equipped with a left-invariant metric, $K \subset G$ be arbitrary subset. Let $\Omega, \Omega' \subset G$ be two subsets such that Ω contains a δ_1- net on K, Ω' contains a δ_2- net on the δ_1- ball $D_{\delta_1} \subset G$. Then the product $\Omega\Omega' \subset G$ contains a δ_2- net on K.*

Proof Take an arbitrary $x \in K$ and some its δ_1- approximant $\omega \in \Omega$. Then $x' = \omega^{-1}x \in D_{\delta_1}$ (the left-invariance of the metric). Take a δ_2- approximant $\omega' \in \Omega'$ of x'. Then $\omega\omega'$ is a δ_2- approximant of x:

$$dist(\omega\omega', x) = dist(\omega', x') < \delta_2.$$

This proves the proposition. □

Let $X > 0$,

$$\varepsilon : \mathbb{R}_+ \to \mathbb{R}_+ \text{ be a decreasing function such that } \varepsilon(cx) < c^{-1}\varepsilon(x) \text{ for any } c > 1,\ x \geq X. \tag{1.1}$$

Example 1.10 For any $\kappa > 0$ the function $\varepsilon(x) = e^{-x^{\kappa}}$ satisfies (1.1) with appropriate X (depending on κ).

Definition 1.11 Let G be a Lie group (equipped with a Riemann metric). Let $(A, B) \in G \times G$ be an irrational pair, $K \subset G$ be a bounded set, $\varepsilon(x)$ be a function as in (1.1). We say that G is $\varepsilon(x)$- *approximable on K by words in* (A, B), if there exist a $c = c(A, B, K) > 0$, a sequence of numbers $l_m = l_m(A, B, K) \in \mathbb{N}$ (called *length majorants*), $l_m \to \infty$, as $m \to \infty$, and a sequence $\Omega_{m,K} = \Omega_{m,K,A,B}$ of word collections such that

$$|w| \leq l_m \text{ for any } w \in \Omega_{m,K} \text{ and} \tag{1.2}$$

$$\text{the subset } \Omega_{m,K}(A, B) \subset G \text{ contains a } \varepsilon(cl_m) - \text{ net on } K. \tag{1.3}$$

We say that G is $\varepsilon(x)$- approximable on K by words in (A, B) *with bounded derivatives*, if $\Omega_{m,K}$ satisfying (1.2) and (1.3) may be chosen so that the union $\cup_m \Omega_{m,K}(A, B)$ is a bounded subset in G and there exist a $\Delta = \Delta(A, B, K) > 0$ and a neighborhood $V \subset G \times G$ of the pair (A, B) such that for any $m \in \mathbb{N}$ and any $w \in \Omega_{m,K}$

$$\text{the mapping } G \times G \to G,\ (a, b) \mapsto w(a, b), \text{ has derivative of norm less than } \Delta \text{ on } V. \tag{1.4}$$

Definition 1.12 We say that a Lie group G is $\varepsilon(x)$- *approximable (with bounded derivatives) by words in* $(A, B) \in G \times G$, if so it is on any bounded subset. We say briefly that G is $\varepsilon(x)$- *approximable (with bounded derivatives)*, if so it is by words in arbitrary irrational pair and on any bounded subset.

The following proposition shows that the $\varepsilon(x)$- approximability is equivalent to the $\varepsilon(x)$-approximability on the unit ball centered at 1.

Proposition 1.13 *Let $\varepsilon(x)$ be as in (1.1), G, (A,B) be as in Definition 1.11, and let the metric on G be left-invariant. Let G be $\varepsilon(x)$- approximable by words in (A,B) (with bounded derivatives) on the unit ball $D_1 \subset G$, $c(A,B,D_1)$, $l_m(D_1) = l_m(A,B,D_1)$, Ω_{m,D_1} be the corresponding constant and sequences of length majorants and word collections, see (1.2) and (1.3). Let $R > 1$, Ω_R be a finite collection of words whose values at (A,B) form a 1- net on D_R,*

$$l(R) = \max_{w \in \Omega_R} |w|.$$

Then G is $\varepsilon(x)$- approximable on D_R by words in (A,B) (with bounded derivatives), where

$$\Omega_{m,D_R} = \Omega_R \Omega_{m,D_1},\ l_m(D_R) = l_m(A,B,D_R) = l(R) + l_m(D_1),\ c(A,B,D_R) = \frac{c(A,B,D_1)}{l(R)}. \tag{1.5}$$

Proof Let Ω_{m,D_R}, $l_m(D_R)$ be the word collections and numbers given by (1.5). For any $m \in \mathbb{N}$ the set $\Omega_{m,D_R}(A,B)$ contains a δ- net on D_R,

$$\delta = \varepsilon(c_1 l_m(D_1)),\ c_1 = c(A,B,D_1),$$

by Proposition 1.9 applied to $K = D_R$, $\Omega = \Omega_R(A,B)$, $\delta_1 = 1$, $\Omega' = \Omega_{m,D_1}(A,B)$, $\delta_2 = \delta$. (The latters satisfy the conditions of the proposition by definition and the $\varepsilon(x)$- approximability.) One has

$$|w| \leq l_m(D_R) \text{ for any } w \in \Omega_{m,D_R},$$

$$\delta \leq \varepsilon(c_1(\inf_m \frac{l_m(D_1)}{l_m(D_R)}) l_m(D_R)) \leq \varepsilon(c(A,B,D_R) l_m(D_R)).$$

This follows by definition, (1.5), the inequality $\frac{l_m(D_1)}{l_m(D_R)} \geq \frac{1}{l(R)}$ and the decreasing of the function $\varepsilon(x)$. If in addition, the set $\cup_m \Omega_{m,D_1}(A,B)$ is bounded and the derivatives of the mappings $(a,b) \mapsto w(a,b)$, $w \in \cup_m \Omega_{m,D_1}$, are uniformly bounded on a neighborhood of (A,B) in $G \times G$, then the same holds true with Ω_{m,D_1} replaced by Ω_{m,D_R} and the same neighborhood. This follows by definition and the finiteness of the collection Ω_R. This proves the $\varepsilon(x)$- approximability on D_R (with bounded derivatives) and Proposition 1.13. □

Corollary 1.14 *Any Lie group $\varepsilon(x)$- approximable by words in a given irrational pair (with bounded derivatives) on unit ball, is $\varepsilon(x)$- approximable by words in the same pair (with bounded derivatives) on any bounded subset.*

The next proposition shows that the notion of $\varepsilon(x)$- approximability is independent on the choice of the metric on G.

Proposition 1.15 *Let $\varepsilon(x)$ be as in (1.1), G, A, B, K be as in Definition 1.11. Let g_1, g_2 be two (complete) Riemann metrics on G. Let the group G equipped with the metric g_1 be $\varepsilon(x)$-approximable on K by words in (A,B) (with bounded derivatives), $\Omega_{m,K}$, $l_m = l_m(A,B,K)$,*

$c_1 = c(A,B,K)$ *be respectively the corresponding word collections, majorants and constant from (1.2) and (1.3). Let*

$$p = \max_m \varepsilon(c_1 l_m), \ K_p \text{ be the closed } p- \text{ neighborhood of } K \text{ in the metric } g_1.$$

Then the group G *equipped with the metric* g_2 *is also* $\varepsilon(x)$*- approximable on* K *by words in* (A,B) *(with bounded derivatives), with respect to the same sequences* $\Omega_{m,K}$*,* l_m *and the new constant*

$$c_2 = c_2(A,B,K) = \rho^{-1}c_1, \ \rho = \max\{\sup_{x,y\in K_p} \frac{d_{g_2}(x,y)}{d_{g_1}(x,y)}, 1\}.$$

Proof Each set $\Omega_{m,K}(A,B)$ contains a $\varepsilon(c_1 l_m)$- net on K in the metric g_1. The latter net is contained in K_p by definition, and is a $\rho\varepsilon(c_1 l_m)$- net on K in the metric g_2 (by the definition of ρ). One has

$$\rho\varepsilon(c_1 l_m) \leq \varepsilon(\rho^{-1}c_1 l_m) = \varepsilon(c_2 l_m), \text{ whenever } m \text{ is large enough,}$$

by definition and (1.1). This proves the $\varepsilon(x)$- approximability in the metric g_2. Let in addition, (G, g_1) (the group G equipped with the metric g_1) be $\varepsilon(x)$- approximable with bounded derivatives, i.e., the set $\cup_m \Omega_{m,K}(A,B)$ be bounded and the derivatives of the mappings $(a,b) \mapsto w(a,b)$, $w \in \cup_m \Omega_{m,K}$, be uniformly bounded on a (bounded) neighborhood $V \subset G \times G$ of (A,B) (in the metric g_1). Then the set $\widetilde{V} = \cup_m \Omega_{m,K}(V)$ is bounded and hence, $\sup_{x,y\in\widetilde{V}} \frac{d_{g_2}(x,y)}{d_{g_1}(x,y)} < +\infty$. The latter inequality together with the previous uniform boundedness of the derivatives on V (in the metric g_1) implies their uniform boundedness on V in the metric g_2. This proves the proposition. □

The following well-known Question is open. It was stated in [17], p.624 (without bounds of derivatives) for the groups $SU(n)$.

Question 1. Is it true that each semisimple Lie group is always $\varepsilon(x)$- approximable with $\varepsilon(x) = e^{-x}$? If yes, does the same hold true with bounded derivatives?

Theorem 1.16 *Let* G *be arbitrary semisimple Lie group (such that there exists at least one irrational pair* $(A,B) \in G \times G$*). Then the group* G *is* $\varepsilon(x)$*- approximable with bounded derivatives, where*

$$\varepsilon(x) = e^{-x^{\kappa}}, \ \kappa = \frac{\ln 1.5}{\ln 9}. \tag{1.6}$$

In addition, for any irrational pair $(A,B) \in G \times G$ *the corresponding length majorants* $l_m = l_m(A,B,D_1)$ *may be chosen so that*

$$l_{m+1} = 9l_m. \tag{1.7}$$

Theorem 1.16 follows from Lemma 1.25 and Theorem 1.26 (both formulated below).

It appears that for many Lie groups the previous approximation rate can be slightly improved. To state the corresponding result, let us introduce the following

Definition 1.17 Let $\mathfrak{g}$ be a Lie algebra with a fixed a positive definite scalar product on it. We say that $\mathfrak{g}$ *has surjective commutator*, if for any $z \in \mathfrak{g} \setminus 0$ there exist $x,y \in \mathfrak{g}$ such that

$$[x,y] = z. \tag{1.8}$$

We say that $\mathfrak{g}$ satisfies *the Solovay-Kitaev inequality*, if there exists a $c > 0$ such that for any $z \in \mathfrak{g} \setminus 0$ there exist $x, y \in \mathfrak{g}$ satisfying (1.8) and such that

$$|x| = |y| < c\sqrt{|z|} \tag{1.9}$$

Theorem 1.18 *(G.Brown, [5]). Each complex semisimple Lie algebra and each real semisimple split Lie algebra (see [21], p.288) have surjective commutator.*

Remark 1.19 In fact, the latter Lie algebras satisfy the Solovay-Kitaev inequality. The author did not find a proof of this statement in the literature, but it can be obtained by minor refinement of Brown's arguments [5]. The question of the surjectivity of commutator in Lie groups has a long history, see [5], [12] and the references therein. We would like to mention one of the first results due to M.Goto [10], who have proved that in any compact semisimple Lie group each element is a commutator of appropriate two other elements.

Example 1.20 The Lie algebras $\mathfrak{su}_n$ satisfy the Solovay-Kitaev inequality [6, 15, 17].

Question 2. Is it true that each real semisimple Lie algebra has surjective commutator? If yes, is it true that it satisfies the Solovay-Kitaev inequality?

Theorem 1.21 *(R.Solovay, A.Kitaev, [6, 15, 17]) Let a Lie group G have a Lie algebra satisfying the Solovay-Kitaev inequality, and there exist at least one irrational pair $(A, B) \in G \times G$. Then the group G is $\varepsilon'(x)$- approximable with*

$$\varepsilon'(x) = e^{-x^{\kappa'}}, \ \kappa' = \frac{\ln 1.5}{\ln 5}. \tag{1.10}$$

In addition, for any irrational pair $(A, B) \in G \times G$ the corresponding length majorants $l_m = l_m(A, B, D_1)$ can be chosen so that

$$l_{m+1} = 5l_m. \tag{1.11}$$

Remark 1.22 In fact, in Theorem 1.21 the Lie group is $\varepsilon'(x)$- approximable with bounded derivatives (with length majorants $l_m(A, B, D_1)$ satisfying (1.11)). This can be easily derived from Kitaev's proof [6, 15, 17]. A proof of this statement is outlined at the end of Subsection 7.2.

Definition 1.23 Let $\mathfrak{g}$ be a Lie algebra with a fixed positive definite scalar product on it. We say that $\mathfrak{g}$ satisfies *the weak Solovay-Kitaev inequality*, if there exists a consant $c > 0$ such that for any $z \in \mathfrak{g} \setminus 0$ there exist $x_j, y_j \in \mathfrak{g}$, $j = 1, 2$, such that

$$z = [x_1, y_1] + [x_2, y_2], \ |x_j| = |y_j| < c\sqrt{|z|}. \tag{1.12}$$

Remark 1.24 The condition that a Lie algebra satisfies a (weak) Solovay-Kitaev inequality is independent on the choice of the scalar product. A Lie algebra satisfying the strong Solovay-Kitaev inequality obviously satisfies the weak one.

Lemma 1.25 *Each semisimple Lie algebra satisfies the weak Solovay-Kitaev inequality.*

Theorem 1.26 *Let a Lie group G have a Lie algebra satisfying the weak Solovay-Kitaev inequality. Let $(A,B) \in G \times G$ be an irrational pair. Then the group G is $\varepsilon(x)$- approximable with bounded derivatives, where $\varepsilon(x)$, $l_m = l_m(A,B,D_1)$ are the same, as in (1.6) and (1.7) respectively.*

Lemma 1.25 is proved in Subsection 7.1. Theorem 1.26 is proved in Subsection 7.2 (analogously to the proof of Theorem 1.21 given in [6, 15, 17]). Together, they imply Theorem 1.16.

1.3 Approximations by groups with relations

Fix a Riemann metric on a Lie group G.

Definition 1.27 Let G be a Lie group, $(A,B) \in G \times G$. Let $\varepsilon(x)$ be a function as in (1.1). We say that the pair (A,B) is $\varepsilon(x)$- *approximable by pairs with relations*, if there exist a $c = c(A,B) > 0$ and sequences of numbers $l_k \in \mathbb{N}$ (called *the length majorants*), $l_k \to \infty$, as $k \to \infty$, nontrivial words $w_k(a,b)$ of lengths at most l_k and pairs $(A_k, B_k) \to (A,B)$ such that for any $k \in \mathbb{N}$ one has

$$w_k(A_k, B_k) = 1 \text{ and } dist((A_k,B_k),(A,B)) < \varepsilon(cl_k) \text{ for any } k \in \mathbb{N}. \tag{1.13}$$

Remark 1.28 The previous Definition and the corresponding word sequence w_k are independent on the choice of the metric on G (while the constant c depends on the metric). The proof of this statement is analogous to the proof of Proposition 1.15.

Theorem 1.29 *Let G be a nonsolvable Lie group, G_{ss} be its semisimple part (see Definition 1.38). Let $\varepsilon(x)$ be a function as in (1.1). Let $A,B \in G$ and $A',B' \in G_{ss}$ be their projections. Let the pair $(A',B') \in G_{ss} \times G_{ss}$ be irrational, and the group G_{ss} be $\varepsilon(x)$- approximable with bounded derivatives by words in (A',B') (see Definition 1.12). Then the pair (A,B) is $\varepsilon(x)$-approximable by pairs with relations.*

Addendum to Theorem 1.29. *In the conditions of Theorem 1.29 the group G_{ss} is $\varepsilon(x)$- approximable by words in (A',B') with bounded derivatives. Let $l_m = l_m(A',B',D_1)$ be the corresponding word length majorants from (1.2). There exist constants $q \in \mathbb{N}$ and $c'' > 0$ depending only on (A,B) such that the pair $(A,B) \in G \times G$ is $\varepsilon(x)$- approximable by pairs with relations having length majorants*

$$l'_m = c'' l_m, \ m \geq q. \tag{1.14}$$

Corollary 1.30 *Each irrational pair of elements in a nonsolvable Lie group is $\varepsilon(x) = e^{-x^\kappa}$-approximable by pairs with relations, where $\kappa = \frac{\ln 1.5}{\ln 9}$, see (1.6). The corresponding length majorant sequence l_k can be chosen so that $l_{k+1} = 9l_k$.*

Proof Let G be a nonsolvable Lie group, $(A,B) \in G \times G$ be an irrational pair. Then its projection $(A',B') \in G_{ss} \times G_{ss}$ is also irrational. The function $\varepsilon(x) = e^{-x^\kappa}$ satisfies the conditions of Theorem 1.29 and its Addendum with a majorant sequence l_k such that $l_{k+1} = 9l_k$ (Theorem 1.16 applied to the semisimple part of G). This together with Theorem 1.29 and its Addendum, see (1.14), implies the corollary. □

Corollary 1.31 *Let G be a nonsolvable Lie group such that the semisimple part of $\mathfrak{g}$ satisfies the Solovay-Kitaev inequality. Then each pair $(A, B) \in G \times G$ with irrational projection to $G_{ss} \times G_{ss}$ is $\varepsilon'(x) = e^{-x^{\kappa'}}$- approximable by pairs with relations, where $\kappa' = \frac{\ln 1.5}{\ln 5}$, see (1.10). The corresponding length majorant sequence l_k can be chosen so that $l_{k+1} = 5l_k$.*

Corollary 1.31 follows from Theorem 1.29 (with the Addendum), Theorem 1.21 and Remark 1.22, analogously to the above proof of Corollary 1.30. Theorem 1.29 together with its Addendum are proved in Section 6.

Question 3. Is it true that in any nonsolvable Lie group each irrational pair of elements is e^{-x}- approximable by pairs with relations?

By Theorem 1.29, a positive solution of Question 1 with bounded derivatives (see 1.2) would imply a positive answer to Question 3.

1.4 Historical remarks and further open questions

The famous Tits' alternative [20] says that any subgroup of linear group satisfies one of the two following incompatible statements:

- either it is solvable up-to-finite, i.e., contains a solvable subgroup of a finite index;
- or it contains a free subgroup with two generators.

Any dense subgroup of a connected semisimple real Lie group satisfies the second statement: it contains a free subgroup with two generators.

The question of possibility to choose the latter free subgroup to be dense was stated in [9] and studied in [4] and [9]. É.Ghys and Y.Carrière [9] have proved the positive answer in a particular case. E.Breuillard and T.Gelander [4] have done it in the general case.

T.Gelander [8] have shown that in any compact nonabelian Lie group any finite tuple of elements can be approximated arbitrarily well by another tuple (of the same number of elements) that generates a nonvirtually free group.

A question (close to Question 1) concerning Diophantine properties of and individual pair $A, B \in SO(3)$ was studied in [14]. We say that a pair $(A, B) \in SO(3) \times SO(3)$ is *Diophantine* (see [14]), if there exists a constant $D > 1$ depending on A and B such that for any word $w_k = w_k(a, b)$ of length k

$$|w_k(A, B) - 1| > D^{-k}.$$

A.Gamburd, D.Jakobson and P.Sarnak have stated the following

Question 4 [7]. *Is it true that almost each pair $(A, B) \in SO(3) \times SO(3)$ is Diophantine?*

V.Kaloshin and I.Rodnianski [14] proved that almost each pair (A, B) satisfies a weaker inequality with the latter right-hand side replaced by D^{-k^2}.

Question 5. *Is there an analogue of Theorem 1.1 for the group of*

- *germs of one-dimensional real diffeomorphisms (at their common fixed point)?*
- *germs of one-dimensional conformal diffeomorphisms?*
- *diffeomorphisms of compact manifold?*

The latter question concerning conformal germs is related to study of one-dimensional holomorphic foliations. A related result was obtained in the joint paper [13] by Yu.S.Ilyashenko and A.S.Pyartli, which deals with one-dimensional holomorphic foliations on $\mathbb{CP}^2$ with isolated singularities and invariant infinity line. They have shown that for a typical foliation

the holonomy group at infinity is free. Here "typical" means "lying outside a zero Lebesgue measure set". It is not known whether this is true for an open set of foliations.

1.5 A simple proof of Theorem 1.1 for $G = PSL_2(\mathbb{R})$

Without loss of generality we assume that $\overline{< A, B >} = G$. Otherwise, $< A, B >$ would be dense in a Lie subgroup of dimension at most two, which is solvable, hence, A and B cannot generate a free subgroup.

The group $G = PSL_2(\mathbb{R})$ acts by conformal transformations of unit disk D_1. There is an open subset $U \subset G$ formed by nontrivial elliptic transformations, which are conformally conjugated to nontrivial rotations. The rotation number (which is the rotation angle divided by 2π) is a local (nowhere zero) analytic function in the parameters of U. An elliptic transformation f has finite order if and only if its rotation number $\rho(f)$ is rational.

Let $w = w(a, b)$ be a word such that $w(A, B) \in U$ (it exists by density). It suffices to show that the function $(a, b) \mapsto \rho(w(a, b))$ is not constant near (A, B): then it follows that there exists a sequence $(a_n, b_n) \to (A, B)$ such that $\rho(w(a_n, b_n)) \in \mathbb{Q}$. Hence, $w(a_n, b_n)$ are finite order elements, thus, one has relations of the type $w^{k_n}(a_n, b_n) = 1$.

The previous function is locally analytic. Suppose the contrary: it is constant. Then by analyticity, it is constant globally and $w(a, b)$ is elliptic with one and the same nonzero rotation number for all the pairs (a, b). On the other hand, it vanishes at $(a, b) = (1, 1)$, since $w(1, 1) = 1$ - a contradiction. This proves Theorem 1.1 for $G = PSL_2(\mathbb{R})$.

1.6 Case of group $Aff_+(\mathbb{R})$.

For any $s > 0$, $u \in \mathbb{R}$ denote

$$g_s : x \mapsto sx, \ t_u : x \mapsto x + u, \ \Gamma(s) =< g_s, t_1 >\subset Aff_+(\mathbb{R}).$$

Proposition 1.32 *For any $s_0 > 0$ there exists a sequence $s_k \to s_0$ such that the corresponding subgroups $\Gamma(s_k)$ have relations that do not hold identically in s.*

Proof It suffices to prove the statement of the proposition for open and dense subset of the values $s_0 > 0$ (afterwards we pass to the closure and diagonal sequences). Thus, without loss of generality we assume that $s_0 \neq 1$. We also assume that $0 < s_0 < 1$, since the groups $\Gamma(s)$ and $\Gamma(s^{-1})$ coincide.

For any s the group $\Gamma(s)$ contains the elements

$$t_{s^k} = g_s^k \circ t_1 \circ g_s^{-k} \text{ and } t_{ms^k}, \ m \in \mathbb{Z}, \ k \in \mathbb{N} \cup 0.$$

We construct sequences of numbers $s_k \to s_0$ and $m_k \in \mathbb{N}$ in such a way that each group $\Gamma(s)$, $s = s_k$, has an extra relation $\tau_{m_k s^k} = \tau_1$. For obvious reasons this is not a relation that holds identically. This will prove the Proposition.

For any k take $m_k = [s_0^{-k}]$, thus, m_k is the integer number such that $m_k s_0^k$ gives a best approximation of 1, with rate less than s_0^k; $m_k s_0^k \to 1$, as $k \to \infty$. The values s_k we are looking for are the positive solutions to the equations $m_k s^k = 1$ (they correspond to the previous relations by definition). Indeed, it suffices to show that $s_k \to s_0$, or equivalently, that the solutions u_k of the equations $\psi_k(u) = m_k(s_0 + u)^k = 1$ converge to 0. The mapping ψ_k is the

composition of the homothety $u \mapsto \widetilde{u} = ku$ and the mapping $\widetilde{\psi}_k : \widetilde{u} \mapsto m_k(s_0 + k^{-1}\widetilde{u})^k$. One has

$$\widetilde{\psi}_k(\widetilde{u}) = m_k s_0^k (1 + k^{-1}\frac{\widetilde{u}}{s_0})^k \to \psi(\widetilde{u}) = e^{\frac{\widetilde{u}}{s_0}}, \text{ as } k \to \infty. \tag{1.15}$$

The convergence is uniform with derivatives on compact sets. The limit $\psi(\widetilde{u})$ is a diffeomorphism $\mathbb{R} \to \mathbb{R}_+$ with unit value at 0. Hence, the solutions $\widetilde{u}_k$ of the equations $\widetilde{\psi}_k(\widetilde{u}) = 1$ converge to 0. Therefore, so do $u_k = k^{-1}\widetilde{u}_k$, $s_k = s_0 + u_k \to s_0$. The proposition is proved. □

1.7 Generalization in the case of semisimple Lie group with irreducible adjoint

Theorem 1.33 *Let G be a semisimple Lie group with irreducible Ad_G (not necessarily connected). Consider a family $\alpha(u) = (a_1(u), \dots, a_M(u))$, $M \in \mathbb{N}$, of M- ples of its elements that depend on a parameter $u \in \mathbb{R}^l$. Let the family $\alpha(u)$ be conj- nondegenerate at 0 (see Definition 1.45 in 1.8). Then there exist arbitrarily small values u such that the mappings $a_i(0) \mapsto a_i(u)$ do not induce group isomorphisms $< \alpha(0) > \to < \alpha(u) >$.*

Theorem 1.33 and Corollary 1.47 (stated below, in 1.8) imply immediately Theorem 1.1 in the case, when G is semisimple, Ad_G is irreducible and A, B generate a dense subgroup. Indeed, suppose the contrary: each pair (a, b) close to (A, B) generates a free subgroup, hence, the mapping $(A, B) \mapsto (a, b)$ induces an isomorphism of the corresponding subgroups. Consider the family of all the pairs (a, b) depending on the parameters in G of the elements a and b. By the previous assumption and Theorem 1.33 (applied to the same family), this family is *conj*- degenerate at (A, B). On the other hand, it is a priori *conj*- nondegenerate at (A, B) (Corollary 1.47), - a contradiction.

1.8 Background material 1: Lie groups, basic definitions and properties.

Everywhere below the Lie algebra of a Lie group G will be denoted

$$\mathfrak{g} = T_1 G.$$

Let us firstly recall what is the adjoint action (see [21], p.32). The group G acts on itself by conjugations (the unity is fixed). The derivative of this action along the vectors of the tangent Lie algebra $\mathfrak{g}$ defines a linear representation of G in $\mathfrak{g}$ called the *adjoint representation*. The adjoint representation of an element $g \in G$ is denoted Ad_g. (If G is a matrix group, then the adjoint action is given by matrix conjugation: $Ad_g(h) = ghg^{-1}$.) The adjoint action of a Lie algebra on itself is defined by the Lie bracket, $\mathrm{ad}_x : y \mapsto [x, y]$. Let G be a Lie group with a given algebra $\mathfrak{g}$. One has

$$Ad_{\exp x} = \exp(\mathrm{ad}_x) \text{ for any } x \in \mathfrak{g}.$$

Definition 1.34 A Lie group is said to be *simple* if the adjoint representation of its unity component is irreducible. A Lie group is said to be *semisimple*, if its unity component has no normal solvable Lie subgroup of positive dimension.

Remark 1.35 A Lie group is (semi)simple, if and only if so is its algebra in the following sense.

Definition 1.36 An *ideal* in a (real or complex) Lie algebra $\mathfrak{g}$ is a Lie subalgebra $I \subset \mathfrak{g}$ (over the corresponding field) such that $[\mathfrak{g}, I] \subset I$. A Lie algebra $\mathfrak{g}$ is said to be *simple*, if it has no nontrivial ideal. A Lie algebra $\mathfrak{g}$ is said to be *semisimple*, if it has no nontrivial *solvable* ideal.

Remark 1.37 A complex Lie algebra is semisimple, if and only if so is it as a real algebra.

It is well-known (see [21], pp. 60, 61) that each Lie algebra $\mathfrak{g}$ has a unique maximal solvable ideal (called *radical*; it may be trivial). The factor of $\mathfrak{g}$ by the radical is a semisimple Lie algebra. Analogously, each nonsolvable Lie group has a unique maximal solvable normal connected Lie subgroup and its tangent algebra coincides with the radical of the Lie algebra of the ambient group; the corresponding Lie group quotient is a semisimple Lie group.

Definition 1.38 The factor of a nonsolvable Lie algebra (group) by its radical (respectively, the maximal solvable normal connected Lie subgroup) is called its *semisimple part*.

Remark 1.39 The Lie algebra of the semisimple part of a nonsolvable Lie group G is the semisimple part of $\mathfrak{g}$.

Remark 1.40 Each semisimple Lie algebra is a finite direct product of simple Lie algebras (the latter product decomposition is unique, see [21], p.151).

Example 1.41 Let $G = SL_n(\mathbb{R})$. The adjoint action of a diagonal matrix

$$A = diag(a_1, \dots, a_n) \in G$$

is diagonalizable and has the eigenvalues 1, $\lambda_{ij} = \frac{a_i}{a_j}$, $i \neq j$. The eigenvector corresponding to the eigenvalue λ_{ij} is represented by the matrix with zeros everywhere except for the (i, j)- th element. The other (unit) eigenvalues correspond to the diagonal matrices. It is well-known that the group $SL_n(\mathbb{R})$ is simple (see, [21], pp. 150, 177).

Proposition 1.42 *For any semisimple (not necessary (simply) connected) Lie group G there exists a collection of semisimple Lie groups $H_1, \dots, H_s$, each one with irreducible adjoint Ad_{H_j}, and a homomorphism*

$$\pi : G \to H_1 \times \cdots \times H_s$$

that is a local diffeomorphism (in particular, $\mathfrak{g} = \prod_{j=1}^{s} \mathfrak{h}_j$). Moreover, the image $\pi(G)$ is projected surjectively onto each group H_j. The kernel of π is contained in the center of the unity component of G.

Proof If the adjoint Ad_G is irreducible, we put $s = 1$, $G = H_1$, and we are done. If G is simply connected, then $\mathfrak{g}$ is a product of simple Lie algebras, and G is the product of the corresponding simply connected Lie groups (which are simple, and hence, have irreducible adjoints).

Case when G is an arbitrary connected semisimple Lie group. Denote $\widetilde{G}$ its universal covering, $C(\widetilde{G})$ the center of $\widetilde{G}$ (which is a discrete subgroup in $\widetilde{G}$). Then

$$G = \widetilde{G}/\Gamma,\ \Gamma \subset C(\widetilde{G}),\ \widetilde{G} = \widetilde{H}_1 \times \cdots \times \widetilde{H}_s,\ \widetilde{H}_j \text{ are simply connected simple groups.}$$

One has $C(\widetilde{G}) = \prod_{j=1}^{s} C(\widetilde{H}_j)$. Therefore, there is a natural projection homomorphism

$$\pi : G = \widetilde{G}/\Gamma \to \widetilde{G}/C(\widetilde{G}) = H_1 \times \cdots \times H_s,\ H_j = \widetilde{H}_j/C(\widetilde{H}_j). \tag{1.16}$$

This is a homomorphism we are looking for.

Case, when G is arbitrary semisimple Lie group. Denote $G_0 \subset G$ its unity component. We assume that Ad_G is not irreducible (the opposite case was already discussed). Let $\mathfrak{g} = \mathfrak{g}_1 \times \cdots \times \mathfrak{g}_r$ be the decomposition of $\mathfrak{g}$ as a product of simple Lie algebras. The adjoint of each $g \in G$ sends any subalgebra $\mathfrak{g}_i$ to an isomorphic subalgebra $\mathfrak{g}_j$; then we say that $\mathfrak{g}_i$ is equivalent to $\mathfrak{g}_j$. To each equivalence class of the $\mathfrak{g}_j$' s we associate the product of the algebras from this class. Denote all the latter products $\mathfrak{h}_1, \dots, \mathfrak{h}_s$: by definition, $\mathfrak{g} = \mathfrak{h}_1 \times \cdots \times \mathfrak{h}_s$. The subalgebras $\mathfrak{h}_j$ are Ad_G- invariant by construction, and $Ad_G|_{\mathfrak{h}_j}$ is irreducible for each j. Indeed, the only Ad_{G_0}- invariant subspaces in $\mathfrak{h}_j$ are the subalgebras $\mathfrak{g}_i$ from the corresponding equivalence class and their products. No one of these subspaces is Ad_G invariant, since Ad_G acts transitively on the subalgebras $\mathfrak{g}_i$ in $\mathfrak{h}_j$ by definition.

Let $\widetilde{H}_j$ be the simply connected Lie groups with algebras $\mathfrak{h}_j$, $\hat{H}_j = \widetilde{H}_j/C(\widetilde{H}_j)$. Let

$$\hat{\pi} : G_0 \to \hat{H}_1 \times \cdots \times \hat{H}_s$$

be the homomorphism (1.16), which is a local diffeomorphism. Consider the subset $H_j' \subset G_0$ of the elements in G_0 whose images under $\hat{\pi}$ have unit $\hat{H}_j$- component: it is the kernel of the composition of $\hat{\pi}$ with the projection to $\hat{H}_j$. This is a normal Lie subgroup in G_0. Denote $H_j^0 \subset H_j'$ its unity component. Its Lie algebra is the product of the $\mathfrak{h}_i$' s with $i \neq j$, which is Ad_G- invariant. Thus, the subgroup $H_j^0 \subset G$ is normal in G. Denote

$$H_j = G/H_j^0;\ \pi : G \to H_1 \times \cdots \times H_s$$

the homomorphism whose components are the natural projections. By construction, this is a local diffeomorphism and the projection of $\pi(G)$ to each H_j is surjective. Denote $\Gamma \subset G$ the kernel of π, which is the intersection of the subgroups $H_j^0 \subset G_0$. It is contained in G_0 and is a discrete normal subgroup there. Hence, it is contained in the center of G_0. Proposition 1.42 is proved. □

Definition 1.43 Let G be a Lie group, $\alpha = (a_1, \dots, a_M) \in G^M$. Consider the G- action on G^M by simultaneous conjugations, $g : \alpha \mapsto g\alpha g^{-1}$, and denote $Conj(a_1, \dots, a_M) \subset G^M$ the orbit of $(a_1, \dots, a_M)$ (i.e., the joint conjugacy class).

Proposition 1.44 *Let G be a semisimple Lie group, $n = dimG$. Let a pair (or M- ple) of its elements be irrational, i.e., generate a dense subgroup in G. Then their joint conjugacy class is bijectively analytically parametrized (as a G- action orbit) by the quotient of the group G by its center. The space of the conjugacy classes corresponding to all the irrational pairs (M- ples) is an analytic manifold of dimension n (respectively, $(M-1)n$). The mapping $(a_1, \dots, a_M) \mapsto Conj(a_1, \dots, a_M)$ is a local submersion at the irrational M- ples $(a_1, \dots, a_M) \in G^M$.*

Proof Let $A = (A_1, \dots, A_M) \in G^M$ be an irrational M- ple: the subgroup $< A >$ generated by A is dense in G. The parametrization $g \mapsto gAg^{-1}$ of the conjugacy class of A by $g \in G$ induces its 1-to-1 parametrization by the quotient of G by its center. Equivalently, for any two distinct elements $g, h \in G$ the elements gAg^{-1}, hAh^{-1} of the conjugacy class of A coincide if and only if $g' = g^{-1}h$ lies in the center of G. Indeed, $gAg^{-1} = hAh^{-1}$, if and only if g' commutes with each A_i, or equivalently, with $< A >$. The latter commutation is equivalent to the commutation with $G = \overline{< A >}$. This proves the previous statement. The irrational M-ples form an open subset in the product of M copies of G (Proposition 1.6). This together with the previous parametrization statement implies the statements of Proposition 1.44. □

Definition 1.45 Let G be a semisimple Lie group, $\alpha(u) = (a_1(u), \dots, a_M(u))$ be a C^1-family of M- ples of its elements depending on a parameter $u \in \mathbb{R}^l$. We say that α is *conj-nondegenerate at* $u = u_0$ if the subgroup $< \alpha(u_0) >\subset G$ is dense in G and the mapping $u \mapsto Conj(\alpha(u))$ has a rank no less than $n = dim G$ at $u = u_0$. Otherwise, if the previous rank is less than n, we say that the family $\alpha(u)$ is *conj*- degenerate at u_0. If $\alpha(u)$ is *conj*-nondegenerate at all u, then we say that $\alpha(u)$ is *conj- nondegenerate.*

Remark 1.46 Let G be a semisimple Lie group, $\alpha(u)$ be an arbitrary family of M- ples of its elements. Then the set of the parameter values u at which $\alpha(u)$ is *conj*- nondegenerate is an open set. This follows from definition and Proposition 1.6.

Corollary 1.47 *Let G be a semisimple Lie group, $(A, B) \in G \times G$ be an irrational pair. The family of all the pairs $(a, b) \in G \times G$ is conj- nondegenerate at (A, B).*

Proof The mapping $(a, b) \mapsto Conj(a, b)$ has full rank at (A, B), which is equal to n (Proposition 1.44). This implies the Corollary. □

For any real linear space (Lie algebra) $\mathfrak{g}$ we denote

$$\mathfrak{g}_{\mathbb{C}} \text{ its complexification,}$$

which is also a linear space (Lie algebra).

1.9 Background material 2: semisimple Lie algebras and root decomposition

Definition 1.48 An element of a Lie algebra is called *regular*, if its adjoint has the minimal possible multiplicity of zero eigenvalue.

Definition 1.49 Let $\mathfrak{g}$ be a complex semisimple Lie group. A *Cartan subalgebra* associated to a regular element of $\mathfrak{g}$ is its centralizer: the set of the elements commuting with it.

It is well-known (see, [21], pp. 153, 159) that
- a) any Cartan subalgebra $\mathfrak{h}$ is a maximal commutative subalgebra;
- b) all the Cartan subalgebras are conjugated;
- c) the adjoint action of $\mathfrak{h}$ on $\mathfrak{g}$ is diagonalizable in an appropriate basis of $\mathfrak{g}$;
- d) the eigenvalues of the latter adjoint action are linear functionals on $\mathfrak{h}$, thus, elements of $\mathfrak{h}^*$, the nonidentically zero ones are called *roots*;

- e) the roots are distinct and the corresponding eigenspaces are complex lines;
- f) if α is a root, then so is $-\alpha$;
- g) for any root α the only roots complex-proportional to α are $\pm\alpha$;
- h) some roots form a complex basis in $\mathfrak{h}^*$ and moreover, an integer root basis in the following sense: each root is an integer linear combination of the basic roots;
- i) the algebra $\mathfrak{g}$ is the direct sum (as a linear space) of $\mathfrak{h}$ and the root eigenlines.

Statement g) follows from the analogous statement in [21] (theorem 6 on p.159) for real-proportional roots and from statement h).

In what follows, for given $\mathfrak{g}$, $\mathfrak{h}$ as above we denote

$$\Delta_{\mathfrak{g}} = \{roots\}, \text{ for any } \alpha \in \Delta_{\mathfrak{g}} \text{ denote } \mathfrak{g}_\alpha \subset \mathfrak{g} \text{ the corresponding eigenline of } \mathrm{ad}_{\mathfrak{h}}. \tag{1.17}$$

For any $\alpha, \beta \in \Delta_{\mathfrak{g}}$ one has

$$[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] = \mathfrak{g}_{\alpha+\beta}, \text{ if } \alpha+\beta \in \Delta_{\mathfrak{g}}; \ [\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subset \mathfrak{h}, \text{ if } \alpha+\beta = 0; \text{ otherwise, } [\mathfrak{g}_\alpha, \mathfrak{g}_\beta] = 0. \tag{1.18}$$

Recall that for any $\alpha \in \Delta_{\mathfrak{g}}$ there exist elements $h_\alpha \in \mathfrak{h}$ (unique) and $e_{\pm\alpha} \in \mathfrak{g}_{\pm\alpha}$ (unique up to multiplication by constants with unit product) such that

$$[e_\alpha, e_{-\alpha}] = h_\alpha, \ [h_\alpha, e_{\pm\alpha}] = \pm 2 e_{\pm\alpha}; \text{ then} \tag{1.19}$$

$$\text{for any root basis } S \subset \Delta_g \text{ the collection } \{h_\alpha\}_{\alpha \in S} \text{ is a basis of } \mathfrak{h} \text{ over } \mathbb{C}. \tag{1.20}$$

Now let $\mathfrak{g}$ be a real semisimple Lie algebra, $x \in \mathfrak{g}$ be a regular element, $\mathfrak{h} \subset \mathfrak{g}$ be the centralizer of x. Then $\mathfrak{h}$ is a maximal commutative subalgebra (which is also called Cartan subalgebra). This follows from the fact that the complexification $\mathfrak{h}_{\mathbb{C}} \subset \mathfrak{g}_{\mathbb{C}}$ is the centralizer of x in $\mathfrak{g}_{\mathbb{C}}$, and hence, is a maximal commutative subalgebra in $\mathfrak{g}_{\mathbb{C}}$ (see statement a) above). Denote $\Delta_{\mathfrak{g}} = \Delta_{\mathfrak{g}_{\mathbb{C}}}$ the collection of the roots of $\mathfrak{h}_{\mathbb{C}}$ in $\mathfrak{g}_{\mathbb{C}}$. The complex conjugation involution $\mathfrak{g}_{\mathbb{C}} \to \mathfrak{g}_{\mathbb{C}}$ has a natural action on the roots. Namely, if a functional $\alpha(x)$ ($x \in \mathfrak{h}_{\mathbb{C}}$) is a root, then

$$\widetilde{\alpha} = \overline{\alpha(\bar{x})} \text{ is also a root. By definition,} \tag{1.21}$$

$$\widetilde{\alpha}(x) = \overline{\alpha(x)} \text{ for any } x \in \mathfrak{h}.$$

We call α a *real (imaginary) root*, if it takes real (imaginary) values on $\mathfrak{h}$. The line $\mathfrak{g}_\alpha$ corresponding to a real root is invariant under the complex conjugation, and its real part $\mathrm{Re}(\mathfrak{g}_\alpha)$ is a real eigenline of $\mathrm{ad}_{\mathfrak{h}}$.

Given a root $\alpha \in \Delta_{\mathfrak{g}}$, let $e_\alpha \in \mathfrak{g}_{\mathbb{C}}$, $h_\alpha \in \mathfrak{h}_{\mathbb{C}}$ be as in (1.19). One has

$$h_{-\alpha} = -h_\alpha, \ h_{\widetilde{\alpha}} = \overline{h}_\alpha, \ \mathfrak{g}_{\widetilde{\alpha}} = \overline{\mathfrak{g}_\alpha}. \tag{1.22}$$

Moreover, one can achieve that $e_{\widetilde{\alpha}} = \pm\overline{e}_\alpha$ by multiplying $e_{\pm\alpha}$ and $e_{\pm\widetilde{\alpha}}$ by appropriate complex constants (even in the case, when α is imaginary, i.e., $\alpha = -\widetilde{\alpha}$; in the opposite case one can always achieve that $e_{\widetilde{\alpha}} = \overline{e}_\alpha$). (We will not use this statement in the paper.) One can choose real e_α for all the real roots α. Denote

$$\Delta^r_{\mathfrak{g}} = \{\text{the real roots in } \Delta_{\mathfrak{g}}\}, \ P_\alpha = \mathrm{Re}(\mathfrak{g}_\alpha) \subset \mathfrak{g} \text{ for any } \alpha \in \Delta^r_{\mathfrak{g}}. \tag{1.23}$$

For each nonreal root α denote

$$\hat{\alpha} = (\alpha, \widetilde{\alpha}) \text{ the nonordered pair of } \alpha \text{ and the conjugated root } \widetilde{\alpha}.$$

To each pair $\hat{\alpha}$ we associate the real 2-plane $P_{\hat{\alpha}} = \mathfrak{g} \cap (\mathfrak{g}_\alpha \oplus \mathfrak{g}_{\widetilde{\alpha}}) = \mathrm{Re}(\mathfrak{g}_\alpha \oplus \mathfrak{g}_{\widetilde{\alpha}})$:

$$P_{\hat{\alpha}} = \{te_\alpha + \bar{t}\bar{e}_\alpha \mid t \in \mathbb{C}\}. \text{ One has } \mathfrak{g} = \mathfrak{h} \oplus E, \ E = (\oplus_{\alpha \in \Delta^r_{\mathfrak{g}}} P_\alpha) \oplus (\oplus_{\hat{\alpha}} P_{\hat{\alpha}}). \tag{1.24}$$

This follows from definition, statement i) (applied to $\mathfrak{g}_{\mathbb{C}}$ and $\mathfrak{h}_{\mathbb{C}}$) and (1.22).

Remark 1.50 The above lines P_α and planes $P_{\hat{\alpha}}$ are $\mathrm{ad}_{\mathfrak{h}}$- invariant.

1.10 Background material 3: proximal elements

Definition 1.51 A linear operator $\mathbb{R}^n \to \mathbb{R}^n$ is called *proximal*, if it has a unique eigenvalue (taken with multiplicity) of maximal modulus (then this eigenvalue is automatically real). An element of a Lie group is proximal, if its adjoint is.

Remark 1.52 The set of proximal operators (elements) is open.

Definition 1.53 A *maximal* $\mathbb{R}$- *split torus* in a semisimple Lie group G is a maximal connected subgroup with a diagonalizable adjoint action on $\mathfrak{g}$ (which is automatically commutative). A semisimple Lie group is called *split* (see [21], p. 288), if some its maximal $\mathbb{R}$- split torus is a maximal connected commutative subgroup.

Example 1.54 Each group $SL_n(\mathbb{R})$ is split: the diagonal matrices form a maximal $\mathbb{R}$- split torus. A typical diagonal matrix is a proximal element of $SL_n(\mathbb{R})$. The group $SO(3)$ is not split, has trivial maximal $\mathbb{R}$- split torus and no proximal elements. The group $SO(2,1)$ is not split and has one-dimensional maximal $\mathbb{R}$- split torus, whose nontrivial elements are proximal in $SO(2,1)$.

Lemma 1.55 *Let a semisimple Lie group contain a proximal element. Then each its maximal* $\mathbb{R}$- *split torus contains a proximal element.*

The proof of Lemma 1.55 is implicitly contained in [1] (p.25, proof of theorem 6.3).

Definition 1.56 An element g of a Lie group will be called *1- proximal,* if the operator $Ad_g - Id$ is proximal.

We use the following equivalent characterization of semisimple Lie groups with proximal elements.

Corollary 1.57 *A semisimple Lie group contains a proximal element, if and only if its unity component contains a 1- proximal element. In this case the 1- proximal elements form an open subset in G accumulating to the unity.*

In the proof of the corollary we use the following properties of the adjoint representation of a semisimple Lie group.

Proposition 1.58 *Let G be a connected semisimple Lie group. For any $x \in \mathfrak{g}$ ($g \in G$) and an eigenvalue λ of* ad_x *(Ad_g) the number $-\lambda$ (respectively, λ^{-1}) is also an eigenvalue of the corresponding adjoint with the same multiplicity, as λ.*

Proof It suffices to prove the statement of the proposition for the Lie algebra: this would imply its statement for any $g \in G$ close enough to 1 (belonging to an exponential chart), and then, for any $g \in G$ (the connectedness of G and the analytic dependence of the operator family Ad_g on $g \in G$). For any regular element $x \in \mathfrak{g}$ the nonzero eigenvalues of ad_x are split into pairs of opposite eigenvalues with equal multiplicities. This follows from the central symmetry of the root system of the complex Cartan subalgebra in $\mathfrak{g}_{\mathbb{C}}$ containing x (see 1.9). The regular elements are dense in $\mathfrak{g}$. This implies that the previous statement remains valid for any $x \in \mathfrak{g}$. This proves the proposition. □

Corollary 1.59 *Any 1- proximal element of a connected semisimple Lie group is proximal.*

Proof Let g be a 1- proximal element, $\lambda \in \mathbb{R}$ be the eigenvalue of $Ad_g - Id$ with maximal modulus (which is simple, and hence, nonzero). Then $(\lambda+1)^{\pm 1}$ are simple eigenvalues of Ad_g (by Proposition 1.58). We claim that $(\lambda+1)^{\pm 1}$ is the eigenvalue of Ad_g with maximal modulus, if $\lambda \in \mathbb{R}_{\pm}$. Indeed, it follows from definition (in both cases) that $(\lambda+1)^{\pm 1} \geq |\lambda|+1$. For any eigenvalue $\lambda' \neq \lambda$ of $Ad_g - Id$ one has $|\lambda'| < |\lambda|$ (1- proximality). This together with the previous and triangle inequalities implies that

$$(\lambda+1)^{\pm 1} \geq |\lambda| + 1 > |\lambda'| + 1 \geq |\lambda' + 1|.$$

This proves the previous statement on the maximality of the eigenvalue $(\lambda+1)^{\pm 1}$ and thus, the proximality of Ad_g. Corollary 1.59 is proved. □

Proposition 1.60 *Let G be a semisimple Lie group, $T \subset G$ be a maximal $\mathbb{R}$- split torus. Let $g \in T$ be a proximal element of G. Then g is also 1- proximal.*

Proof The eigenvalues of Ad_g (which are real, since $Ad_T : \mathfrak{g} \to \mathfrak{g}$ is diagonalizable) are positive, since this is true for $Ad_1 = Id$ and the torus T is connected. The nonunit eigenvalues are split into pairs of inverses (Proposition 1.58). Hence, we can order them as follows (distinct indices correspond to distinct (may be multiple) eigenvalues):

$$0 < \lambda_1^{-1} < \lambda_2^{-1} < \cdots < \lambda_k^{-1} < 1 < \lambda_k < \cdots < \lambda_1. \tag{1.25}$$

The eigenvalue λ_1 is simple (proximality). One has

$$\lambda_1 - 1 > \lambda_1^{-1}(\lambda_1 - 1) = 1 - \lambda_1^{-1}, \text{ since } 0 < \lambda_1^{-1} < 1$$

by (1.25). This together with (1.25) implies that $\lambda_1 - 1$ is a simple eigenvalue of $Ad_g - Id$ with maximal modulus. Hence, the operator $Ad_g - Id$ is proximal. Proposition 1.60 is proved. □

Proof of Corollary 1.57. Let the unity component of G contain a 1- proximal element. Then this element is proximal (Corollary 1.59). Conversely, let G contain proximal elements. Let $T \subset G$ be a maximal $\mathbb{R}$- split torus, $g \in T$ be a proximal element of G (which exists by Lemma 1.55). Then g is 1- proximal (Proposition 1.60) and lies in the unity component of G.

Now let us prove the last statement of Corollary 1.57. To do this, consider the 1- parameter subgroup $\Gamma \subset T$ passing through the previous proximal element g. The set $\Gamma \setminus 1$ consists of proximal elements, since Ad_g is proximal and any positive power of a proximal operator is also proximal. Therefore, $\Gamma \setminus 1$ consists of 1- proximal elements (Proposition 1.60). The 1- proximal elements in G form an open subset (Remark 1.52) that accumulates to 1 (at least along the group Γ). This proves the corollary. □

2 Proof of Theorems 1.1 and 1.33 for semisimple Lie groups with irreducible adjoint and proximal elements

Here and in Section 3 we prove Theorem 1.33, which deals with semisimple Lie groups having irreducible adjoint representation. For those Lie groups Theorem 1.1 follows from Theorem 1.33 (see 1.7). In the present section we treate the case of Lie group with proximal elements. The opposite case is treated in the next section.

2.1 Motivation and the plan of the proof

Let G be a semisimple Lie group with irreducible adjoint and proximal elements, $n = dimG$, $\alpha(u) = (a_1, \dots, a_M)(u)$ be a *conj*- nondegenerate at $u = 0$ family of M- ples of elements of G depending on parameter u (see Definition 1.45). Recall that the subgroup $< \alpha(0) > \subset G$ is dense. Without loss of generality we assume that

- the parameter space has the same dimension n, as G: $u \in \mathbb{R}^n$ (we can restrict our family to appropriate generic n- dimensional subspace in the parameter space, along which the family remains *conj*- nondegenerate).

We construct a sequence w_k of words in M elements such that there exists a sequence $u_k \in \mathbb{R}^n$ for which

$$w_k(\alpha(u_k)) = 1, \ u_k \to 0, \ \text{as } k \to \infty, \tag{2.1}$$

and the relations $w_k(\alpha(u)) = 1$ do not hold true identically in a neighborhood of 0. Then the mapping $\alpha(0) \mapsto \alpha(u)$ does not extend up to a group isomorphism $< \alpha(0) > \to < \alpha(u) >$ for arbitrarily small values of u. Indeed, the relations $w_k = 1$ hold true in the group $< \alpha(u) >$ for the values $u = u_k$ (which tend to 0), and do not hold for some other values of u (which can be chosen arbitrarily small as well). This will prove Theorem 1.33.

Firstly let us motivate the proof of Theorem 1.33. A natural way to construct the previously mentioned words w_k is to achieve that $w_k(\alpha(0)) \to 1$. Then to guarantee the existence of a sequence $u_k \to 0$ of solutions to the equations $w_k(\alpha(u)) = 1$, we have to show that there exists a sequence $\delta_k \to 0$ such that $1 \in w_k(\alpha(D_{\delta_k}))$, whenever k is large enough. To do this, we have to prove an appropriate lower bound for derivatives of the mappings $w_k(\alpha(u))$ near 0; in particular, to show that certain derivatives will be greater than $\delta_k^{-1} dist(w_k(\alpha(0)), 1)$.

By density, we can always construct a sequence of words w_k so that $w_k(\alpha(0)) \to 1$. In the case, when $a_i(0)$ are close enough to unity, it suffices to take w_k to be a sequence of appropriate successive commutators

$$w_1 = []_1 = [a_1, a_2], \ w_2 = []_2 = [a_1, [a_1, a_2]], \dots$$

On the other hand, the derivatives of the corresponding mappings $w_k(\alpha(u))$ do not admit a satisfactory lower bound: the values at $\alpha(0)$ of the commutators converge exponentially to 1, and the previous derivatives (taken at 0) converge exponentially to zero.

In order to construct words w_k with large derivatives, we use the following observation. Fix a small $\Delta > 0$. Then $dist([]_k(\alpha(0)), 1) < \Delta$, whenever k is large enough. Consider all the powers $[]_k^m$ of the previous commutators. Put

$$m_k = \min\{m \in \mathbb{N},\ dist([]_k^m(\alpha(0)), 1) \geq \Delta\}.$$

(The numbers m_k are well-defined provided that $[]_k(\alpha(0)) \neq 1$, which holds true "generically".) Then $\Delta \leq dist([]_k^{m_k}(\alpha(0)), 1) < 2\Delta$, whenever k is large enough, by definition and the previous inequality. We claim that if $a_1(0)$ and $a_2(0)$ are close enough to 1 and satisfy appropriate genericity assumption, then the derivative at 0 in certain directions of the mapping $u \mapsto []_k^{m_k}(\alpha(u)) \in G$ grows linearly in k, as that of the mappings ψ_k in the proof of Proposition 1.32.

In what follows we construct

- appropriate words $g_1, \dots, g_n$, h, w and define recurrently the iterated commutators

$$w_{i0} = h,\ w_{ik} = g_i w_{i(k-1)} g_i^{-1} w_{i(k-1)}^{-1}, \tag{2.2}$$

- a sequence of collections

$$M_k = (m_{1k}, \dots, m_{nk}),\ m_{ik} \in \mathbb{N},\ \text{and put}$$

$$\omega_k = w_{1k}^{m_{1k}} \dots w_{nk}^{m_{nk}},\ w_k = w^{-1}\omega_k. \tag{2.3}$$

We show that the latter words w_k satisfy (2.1). To do this, we introduce the rescaled parameter

$$\widetilde{u} = ku,$$

as in Proposition 1.32, and show that

$$\omega_k(\alpha(k^{-1}\widetilde{u})) \to \Psi(\widetilde{u}),\ \text{as } k \to \infty;\ \Psi : \mathbb{R}^n \to G \text{ is a local diffeomorphism at } 0, \tag{2.4}$$

the previous convergence is uniform with derivatives on compact subsets in $\mathbb{R}^n$. Theorem 1.33 will be then deduced from (2.4) at the end of the subsection.

For a fixed $g \in G$ consider the corresponding commutator mapping

$$\phi_g : G \to G,\ \phi_g(h) = ghg^{-1}h^{-1}. \text{ One has } \phi_g(1) = 1,\ \phi_g'(1) = Ad_g - Id : \mathfrak{g} \to \mathfrak{g},$$

$$w_{ik}(\alpha(u)) = \phi^k_{g_i(\alpha(u))}(h(\alpha(u))). \tag{2.5}$$

For any 1- proximal element $g \in G$ (see Definition 1.56) denote

$$s(g) = \text{ the eigenvalue of } Ad_g - Id \text{ with maximal modulus, } L_g \subset \mathfrak{g} \text{ its eigenline.} \tag{2.6}$$

The function $s(g)$ is analytic on the (open) subset of 1- proximal elements, by the simplicity of the eigenvalue $s(g)$. Denote

$$\Pi = \{1\text{- proximal elements } g \in G \mid |s(g)| < 1\}, \tag{2.7}$$

Remark 2.1 Let G be an arbitrary semisimple Lie group with proximal elements. The above set Π is open and nonempty (Corollary 1.57).

The choice of the words g_j and h will be specified at the end of the subsection. It will be done so that

$$g_j(\alpha(0)) \in \Pi \text{ for any } j = 1, \dots, n.$$

The following Proposition 2.2 (proved in 2.2) describes the asymptotic behavior of the iterated commutators $\phi_g^k(h)$, as $k \to \infty$, for arbitrary $g \in \Pi$. Using Proposition 2.2, we show (Corollary 2.3) that for appropriately chosen word h and arbitrary given $\varepsilon > 0$ one can choose appropriate exponents m_{jk} (which depend on g_j and ε, see (2.11)) so that the mapping sequence $\omega_k(\alpha(k^{-1}\widetilde{u}))$ converges to some mapping $\Psi(\widetilde{u})$, which depends only on g_j, h and ε. The mapping Ψ is explicitly given by formula (2.12) below. The main technical part of the proof of Theorem 1.33 is to show that one can adjust g_j, h and ε so that the limit Ψ be a local diffeomorphism at 0 (Lemmas 2.4, 2.6 and the Main Technical Lemma 2.5 below). Lemma 2.5 is proved in 2.4. Lemmas 2.4 and 2.6 are deduced from it in the present subsection and in 2.3 respectively. Theorem 1.33 will be deduced from Lemma 2.6 and Proposition 2.2 at the end of the subsection.

Proposition 2.2 *Let G be a Lie group with proximal elements, Π be as in (2.7). There exist an open subset*

$$\Pi' \subset \Pi \times G, \ \Pi' \supset \Pi \times 1, \tag{2.8}$$

and a $\mathfrak{g}$- valued vector function $v_g(h)$ analytic in $(g, h) \in \Pi'$ (denote $dv_g : \mathfrak{g} \to \mathfrak{g}$ its differential in h at $h = 1$) such that $v_g(1) = 0$ and for any $(g, h) \in \Pi'$ one has

$$v_g(h) \in L_g, \ dv_g|_{L_g} = Id : L_g \to L_g, \ \phi_g^k(h) = \exp(s^k(g)(v_g(h) + o(1))), \ \text{as } k \to +\infty, \tag{2.9}$$

see Fig.1; $s(g)$ and L_g are the same, as in (2.6). The latter "o" is uniform with derivatives in (g, h) on compact subsets in Π'.

The proposition is proved in Subsection 2.2.

Corollary 2.3 *Let G, n, M, $\alpha(u)$ be as at the beginning of the subsection, Π be as in (2.7), Π', v_g be as in Proposition 2.2. Let $g_1, \dots, g_n$, h be words in M elements such that*

$$(g_j(\alpha(0)), h(\alpha(0))) \in \Pi' \text{ for any } j = 1, \dots, n. \text{ Put}$$

$$s_j(u) = s(g_j(\alpha(u))), \ \widetilde{\nu}_j(u) = v_{g_j(\alpha(u))}(h(\alpha(u))) \in \mathfrak{g}, \ \nu_j = \widetilde{\nu}_j(0). \tag{2.10}$$

Let $\varepsilon > 0$. For any $k \in \mathbb{N}$ and $j = 1, \dots, n$ put

$$m_{jk} = [\varepsilon |s_j|^{-k}(0)]. \tag{2.11}$$

Let ω_k be the corresponding commutator power product (2.3). Then

$$\omega_k(\alpha(k^{-1}\widetilde{u})) \to \Psi(\widetilde{u}) = \exp(\varepsilon e^{(d \ln s_1(0))\widetilde{u}} \nu_1) \dots \exp(\varepsilon e^{(d \ln s_n(0))\widetilde{u}} \nu_n), \ \text{as } k \to \infty, \tag{2.12}$$

uniformly with derivatives on compact subsets in $\mathbb{R}^n$.

Proof One has

$$w_{jk}^{m_{jk}}(\alpha(k^{-1}\widetilde{u})) \to \exp(\varepsilon e^{(d\ln s_j(0))\widetilde{u}}\nu_j) \tag{2.13}$$

uniformly with derivatives on compact sets in $\mathbb{R}^n$. Indeed, by (2.5) and (2.9), one has

$$w_{jk}^{m_{jk}}(\alpha(k^{-1}\widetilde{u})) = \exp(m_{jk}s_j^k(k^{-1}\widetilde{u})(\widetilde{\nu}_j(k^{-1}\widetilde{u}) + o(1))), \ \widetilde{\nu}_j(k^{-1}\widetilde{u}) \to \nu_j, \tag{2.14}$$

$$m_{jk}s_j^k(k^{-1}\widetilde{u}) \to \varepsilon e^{(d\ln s_j(0))\widetilde{u}}, \text{ since} \tag{2.15}$$

$$s_j^k(k^{-1}\widetilde{u}) = (s_j(0) + k^{-1}(ds_j(0))\widetilde{u} + o(k^{-1}))^k$$

$$= s_j^k(0)(1 + k^{-1}(d\ln s_j(0))\widetilde{u} + o(k^{-1}))^k = s_j^k(0)e^{(d\ln s_j(0))\widetilde{u}}(1 + o(1)) \tag{2.16}$$

and $m_{jk}s_j^k(0) \to \varepsilon$ by (2.11). Substituting (2.15) to (2.14) yields (2.13), which implies (2.12). The corollary is proved. □

Lemma 2.4 *Let G, n, $\alpha(u)$, M be as at the beginning of the subsection, Π be as in (2.7). There exists a collection $g_1, \dots, g_n$ of words in M elements such that $g_i(\alpha(0)) \in \Pi$ for all $i = 1, \dots, n$ and the system of n functions $s_i(u) = s(g_i(\alpha(u)))$ (which are well-defined in a neighborhood of 0) has the maximal rank n at 0. Moreover, given any collection $A_1, \dots, A_n \in \Pi$ one can achieve that in addition, the elements $g_i(\alpha(0))$ be arbitrarily close to A_i.*

For the proof of Theorem 1.33 in the general case, without the assumption that G has proximal elements, we use the following generalization of Lemma 2.4.

Lemma 2.5 (Main Technical Lemma). *Let G be an arbitrary semisimple Lie group with irreducible adjoint representation (not necessarily with proximal elements), $dimG = n$. Let $\alpha(u) = (a_1(u), \dots, a_M(u))$ be a conj- nondegenerate at 0 family of M- ples of its elements depending on a parameter $u \in \mathbb{R}^n$. Let $U \subset G$ be an arbitrary open subset and $\sigma : U \to \mathbb{R}$ be a smooth locally nonconstant function. Then there exist n abstract words $g_i(a_1, \dots, a_M)$, $i = 1, \dots, n$, such that the system of n functions $s_i(u) = \sigma(g_i(\alpha(u)))$ is well-defined (locally near 0) and has the maximal rank n at 0. Moreover, for any given $A_1, \dots, A_n \in U$ one can achieve that in addition, the elements $g_i(\alpha(0)) \in G$ be arbitrarily close to A_i.*

Lemma 2.5 will be proved in 2.4. Lemma 2.4 follows from Lemma 2.5 applied to $U = \Pi$ and the function $\sigma(g) = s(g)$.

Lemma 2.6 *Let G, n, M, $\alpha(u)$ be as at the beginning of the subsection, $\Pi' \subset G \times G$ be as in Proposition 2.2. There exist words g_j, h, $j = 1, \dots, n$, such that $(g_j(\alpha(0)), h(\alpha(0))) \in \Pi'$ for all j and for any $\varepsilon > 0$ small enough the corresponding mapping $\Psi(\widetilde{u})$ from (2.12) is a local diffeomorphism at 0.*

Lemma 2.6 will be deduced from Lemma 2.4 in Subsection 2.3.

Proof of Theorem 1.33 modulo Proposition 2.2 and Lemmas 2.5 and 2.6. Let g_j, h, ε be as in Lemma 2.6, $s_j(u)$ be as in (2.10), m_{jk} be as in (2.11). Let ω_k be the corresponding commutator power product from (2.3), Ψ be the mapping from (2.12). Let $\delta > 0$ be such that $\Psi : \overline{D}_\delta \to \Psi(\overline{D}_\delta) \subset G$ be a diffeomorphism (it exists by Lemma 2.6). Let w be an arbitrary word such that

$$w(\alpha(0)) \in \Psi(D_\delta), \ w_k = w^{-1}\omega_k. \text{ Then}$$

$$w_k(\alpha(k^{-1}\widetilde{u})) \to \psi(\widetilde{u}) = w^{-1}(\alpha(0))\Psi(\widetilde{u}),\ \psi: \overline{D}_\delta \to \psi(\overline{D}_\delta) \subset G \text{ is a diffeomorphism}, \tag{2.17}$$
$$1 \in \psi(D_\delta)$$

(Corollary 2.3). Therefore, for any k large enough the image $w_k(\alpha(k^{-1}D_\delta))$ also contains 1, and hence, $w_k(\alpha(k^{-1}\widetilde{u}_k)) = 1$ for some $\widetilde{u}_k \in D_\delta$. Put

$$u_k = k^{-1}\widetilde{u}_k;\ \text{one has } w_k(\alpha(u_k)) = 1,\ u_k \to 0.$$

The relation $w_k(\alpha(u)) = 1$, which holds for $u = u_k$, does not hold identically in u, because of the diffeomorphicity of the mappings $\widetilde{u} \mapsto w_k(\alpha(k^{-1}\widetilde{u}))$ on D_δ for large k (see (2.17); the convergence is uniform with derivatives on $\overline{D}_\delta$ there). Thus, the words w_k satisfy (2.1). This proves Theorem 1.33. □

2.2 Dynamics of commutator. Proof of Proposition 2.2

The commutator mapping $\phi = \phi_g$ corresponding to $g \in \Pi$ represents a germ of analytic mapping $(G, 1) \to (G, 1)$ at its fixed point 1 that has the following property:

$$\text{its derivative at the fixed point has a simple real eigenvalue } s = s(g),\ 0 < |s| < 1, \tag{2.18}$$

that is greater than the modulus of its any other complex eigenvalue.

This property implies the following dynamical corollaries.

Proposition 2.7 *Let $\phi : (\mathbb{R}^n, 0) \to (\mathbb{R}^n, 0)$ be a germ of analytic mapping at its fixed point 0 satisfying (2.18), s be the corresponding eigenvalue of the differential $D\phi(0)$. There exist a neighborhood of zero in $\mathbb{R}^n$ and local analytic coordinates (x_1, x_2) on this neighborhood, $x_1 \in \mathbb{R}$, $x_2 \in \mathbb{R}^{n-1}$, where the mapping takes the form*

$$\phi : (x_1, x_2) \mapsto (sx_1, Q(x_1, x_2)),\ \frac{\partial Q}{\partial x_1}(0, 0) = 0,\ ||Q(x)|| < \lambda||x||,\ \lambda = const,\ 0 < \lambda < s. \tag{2.19}$$

If ϕ depends (real-) analytically on some parameter, then the corresponding coordinates (x_1, x_2) may be chosen to depend analytically on the same parameter. The inequality in (2.19) remains valid (locally in the parameter) with one and the same λ.

The proposition follows from a version of Poincaré-Dulac theorem ([2], chapter 5, section 25, subsection D). In more detail, firstly we reduce ϕ by a linear variable change to the form $\phi : (x_1, x_2) \mapsto (sx_1 + o(|x_1| + |x_2|), Q(x_1, x_2))$. Then one can kill the terms $o(|x_1| + |x_2|)$ (which are nonresonant by (2.18)) by analytic variable change, by Poincaré-Dulac theorem (the definition of nonresonant terms may be found in loc. cit).

Remark 2.8 In the previous proposition the mapping ϕ always has a local invariant analytic hypersurface $x_1 = 0$. In particular, so does the commutator mapping ϕ_g, whenever $g \in \Pi$; the corresponding local invariant hypersurface through 1 (we denote it S_g) is transversal to L_g, see Fig.1. In general, the hypersurface S_g is not a Lie subgroup. Indeed, consider the case, when $G = SL_3(\mathbb{R})$ and g is a generic diagonal matrix (as in Example 1.41) close enough to 1. Then the tangent hyperplane $T_1 S_g$ is spanned by the diagonal matrices and the standard

one-element nilpotent matrices, except for the one-element matrix $v'_g \in \mathfrak{sl}_3$ corresponding to the maximal eigenvalue of $Ad_g - Id$. The hyperplane $T_1 S_g$ is not a Lie subalgebra: a nonzero vector $v'_g \in L_g$ is a commutator of two appropriate nilpotent matrices in $T_1 S_g$.

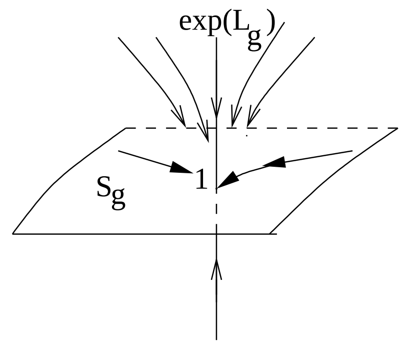

Figure 1: The dynamics of ϕ_g

Proof of Proposition 2.2. The germs of the commutator mappings ϕ_g, $g \in \Pi$, at their fixed point 1 satisfy the conditions of Proposition 2.7 and depend analytically on the parameter $g \in \Pi$. Let $(x_{1,g}, x_{2,g})$ be the germs at 1 of the corresponding coordinates from (2.19). There exists a neighborhood $\Pi' \subset \Pi \times G$ of the unit section $\Pi \times 1$ such that $x_{1,g}$, $x_{2,g}$ may be chosen to depend analytically on $(g, h) \in \Pi'$. Without loss of generality we consider that for any $g \in \Pi$ the domain $\{h \in G \mid (g, h) \in \Pi'\}$ is mapped into itself by ϕ_g, and the corresponding numbers $\lambda = \lambda(g) < s(g)$ from (2.19) be uniformly bounded on compact subsets in Π. One can achieve this by shrinking Π' in an appropriate way. Then for any $(g, h) \in \Pi'$ one has

$$\phi_g^k(h) = (s^k(g) x_{1,g}(h), 0) + o(s^k(g)), \text{ as } k \to \infty \tag{2.20}$$

in the coordinates $(x_{1,g}, x_{2,g})$; the latter "o" being uniform with derivatives on compact sets in Π'. This follows from (2.19). The tangent vector

$$\nu_g = \frac{\partial}{\partial x_{1,g}} \in \mathfrak{g} = T_1 G \text{ is contained in the line } L_g,$$

by definition and since ν_g is an eigenvector of the differential $d\phi_g(1)$ with the eigenvalue $s(g)$, see (2.19). Put

$$v_g(h) = x_{1,g}(h) \nu_g \in L_g \text{ for any } (g, h) \in \Pi'.$$

Let us prove statements (2.9) for this vector function $v_g(h)$.

The exponential mapping $\exp : \mathfrak{g} \to G$ has unit derivative $Id : \mathfrak{g} \to \mathfrak{g}$ at 0. The mapping of the $x_{1,g}$- axis in G to L_g given in the coordinates by $\Lambda_g : (\tau, 0) \mapsto \tau \nu_g \in L_g$ has unit derivative $Id : L_g \to L_g$ at $(0, 0) = 1 \in G$ (by the definition of the vector ν_g). Therefore, the composition $\exp \circ \Lambda_g$ has unit derivative $Id : L_g \to L_g$. In particular,

$$(\tau, 0) = \exp(\tau \nu_g + o(\tau)), \text{ as } \tau \to 0.$$

The right-hand side in (2.20) equals $\exp(s^k(g) x_{1,g}(h) \nu_g + o(s^k(g))) = \exp(s^k(g) v_g(h) + o(s^k(g)))$, as $k \to \infty$. This follows from definition and the previous formula applied to

$\tau = s^k(g)x_{1,g}(h)$. This together with (2.20) proves the asymptotic formula in (2.9). The differential at $h = 1$ of the vector function $v_g(h)$ equals $dv_g = \nu_g dx_{1,g}$. The restriction to L_g of the latter differential is the identity operator $Id : L_g \to L_g$ by definition. This proves (2.9) and Proposition 2.2. □

2.3 The diffeomorphicity of Ψ. Proof of Lemma 2.6 modulo the Main Technical Lemma

Let g_j be some words such that

$$g_j(\alpha(0)) \in \Pi \text{ for all } j;\ s_j(u) = s(g_j(\alpha(u))). \text{ Denote } \sigma_j = d\ln s_j(0) : \mathbb{R}^n \to \mathbb{R}.$$

We use the following sufficient condition for the local diffeomorphicity of the mapping Ψ at 0.

Proposition 2.9 *Let G be a Lie group, $n = dimG$. Let $\sigma_1, \dots, \sigma_n : \mathbb{R}^n \to \mathbb{R}$ be a collection of linearly independent 1- forms, $\nu_1, \dots, \nu_n \in \mathbb{R}^n$ be a collection of linearly independent vectors. Then for any $\varepsilon > 0$ small enough*

$$\Psi(\widetilde{u}) = \exp(\varepsilon e^{\sigma_1(\widetilde{u})}\nu_1)\dots\exp(\varepsilon e^{\sigma_n(\widetilde{u})}\nu_n) \textit{ is a local diffeomorphism at } 0.$$

Proof Denote

$$\widetilde{\psi}(\widetilde{u}) = (\Psi(0))^{-1}\Psi(\widetilde{u}). \text{ By definition, } \widetilde{\psi}(0) = 1.$$

The derivative $\widetilde{\psi}'(0)$ is a linear operator $T_0\mathbb{R}^n \to \mathfrak{g}$ that is $O(\varepsilon^2)$- close to

$$\Omega = \varepsilon \sum_{j=1}^{n} \nu_j \sigma_j : T_0\mathbb{R}^n \to \mathfrak{g},$$

as $\varepsilon \to 0$. The latter sum is a nondegenerate operator (the linear independence of σ_j and ν_j). The two previous statements together imply the proposition. □

As it is shown below, Lemma 2.6 is implied by Lemma 2.5, Proposition 2.9 and the following

Proposition 2.10 *Let G be a Lie group with irreducible adjoint and proximal elements, $\Pi \subset G$ be as in (2.7). Let $0 < s < 1$ be a value such that $s = s(g)$ for some $g \in \Pi$. The lines $L_g \subset \mathfrak{g}$ (see (2.6)) with $g \in \{s(g) = s\} \subset \Pi$ generate the whole linear space $\mathfrak{g}$.*

Proof The subspace in $\mathfrak{g}$ generated by the previous lines is nonzero and Ad_G- invariant (by definition and since the set Π is invariant under conjugations). This together with the irreducibility of Ad_G proves the proposition. □

Choice of the words g_j. We use Lemma 2.4, which follows from Lemma 2.5, see 2.1. Let $A_1, \dots, A_n \in \Pi$ be such that

$$s = s(A_1) = \dots = s(A_n);\ \text{the lines } L_{A_1}, \dots, L_{A_n} \text{ are linearly independent}$$

(their existence follows from Proposition 2.10). Fix words g_j from Lemma 2.4 so that

$$A'_j = g_j(\alpha(0)) \in \Pi,\ \text{the lines } L_{A'_j} \text{ are linearly independent:}$$

it is sufficient to achieve that A'_j be close enough to A_j (Lemma 2.4).

Choice of the word h. Let us choose h so that

$$(A'_j, h(\alpha(0))) \in \Pi',\ \nu_j = v_{A'_j}(h(\alpha(0))) \neq 0 \text{ for any } j = 1, \dots, n :$$

it is sufficient to achieve that $h(\alpha(0))$ be close enough to 1 and do not lie in the hypersurfaces $v_{A'_j} = 0$ (the inequality $dv_{A'_j} \neq 0$, see (2.9), and the density of the subgroup $< \alpha(0) >\subset G$).

The 1- forms $\sigma_j = d \ln s_j(0)$ and the vectors ν_j are linearly independent by Lemma 2.4 and construction. The corresponding mapping Ψ is a local diffeomorphism at 0, whenever ε is small enough (Proposition 2.9). This proves Lemma 2.6 modulo the Main Technical Lemma 2.5. Theorem 1.33 is proved modulo the Main Technical Lemma.

2.4 Independent eigenvalues. Proof of Main Technical Lemma 2.5

Denote $\hat{U} = \mathbb{R}^n$ the parameter u space under consideration. By assumption, the family $\alpha(u)$ is *conj*- nondegenerate. This together with the equality of the dimensions of G and $\hat{U}$ implies that the derivative along each nonzero vector $v \in T_0\hat{U}$ of the function $u \mapsto Conj(\alpha(u))$ is nonzero. (Fix a $v \in T_0\hat{U} \setminus 0$.) The derivatives along v of the mappings $u \mapsto w(\alpha(u))$ (where w is an arbitrary word) form a vector field on the dense subgroup $\Gamma =< \alpha(0) >\subset G$ (we extend it to 1 by 0). This vector field is well-defined (single-valued), if Γ is free. In general, if there are relations in Γ, it is single-valued, if and only if for any word w giving a relation (i.e., $w(\alpha(0)) = 1$) the corresponding mapping $u \mapsto w(\alpha(u))$ has zero derivative along v.

The maximal rank statement of Lemma 2.5 is equivalent to the statement that for any given $v \in T_0\hat{U} \setminus 0$ there exists an index j such that the corresponding vector at $g_j(\alpha(0))$ of the previous field is nonzero and transversal to the level hypersurface of the function σ. To prove that, we show (in the next Lemma 2.11) that the previous vector field (if well-defined) is not Lipschitz at 1. Namely, we show (in the proof of Lemma 2.11) that if it were Lipschitz, it would define an infinitesimal automorphism of G (hence, this is an interior automorphism, thus preserving conjugacy classes). This would contradict the nonvanishing of the derivative of $Conj(\alpha(u))$. We also show that the lines generated by those derivatives that are large with respect to $dist(w(\alpha(0)), 1)$ approach any given line in $\mathfrak{g}$, as $w(\alpha(0)) \to 1$. The previous transversality statement will be then deduced in the next corollary.

Lemma 2.11 *Let G be a semisimple Lie group with irreducible Ad_G, $\alpha(u) = (a_1, \dots, a_M)(u)$ be a smooth family of M- ples of its elements depending on a parameter $u \in \hat{U} = \mathbb{R}^n$, $n = dimG$. Let the subgroup $< \alpha(0) >\subset G$ be dense. Let $v \in T_0\hat{U} \setminus 0$ be such that the derivative along v of the function $u \mapsto Conj(\alpha(u))$ does not vanish. Then for any line $\Lambda \subset \mathfrak{g}$, $0 \in \Lambda$, there exists a sequence of words $w_k(a_1, \dots, a_M)$, $w_k(\alpha(0)) \to 1$, as $k \to \infty$, with the following properties:*

1) Consider the derivatives of the functions $w_k(\alpha(u))$ along v as a vector field at the points $\{w_k(\alpha(0))\}$. This field is not Lipschitz at 1, more precisely,

$$\frac{|\frac{dw_k(\alpha(u))}{dv}|}{dist(w_k(\alpha(0)), 1)} \to \infty, \text{ as } k \to \infty. \tag{2.21}$$

2) The tangent line to G at $w_k(\alpha(0))$ generated by the latter derivative tends to Λ.

Corollary 2.12 *Let G, n, $\alpha(u)$, $U\subset G$, $\sigma: U\to\mathbb{R}$ be the same, as in the Main Technical Lemma 2.5. Let $v\in T_0\mathbb{R}^n$, $v\neq 0$. For any $g\in U$ there exists a sequence of words $\widetilde{w}_k$, $h_k=\widetilde{w}_k(\alpha(0))\to g$, such that the derivatives $\frac{d\widetilde{w}_k(\alpha(u))}{dv}$ are transversal to the level hypersurface $\sigma=\sigma(h_k)$.*

The lemma and its corollary are proved below.

Proof of Lemma 2.5. Given a $\varepsilon>0$ and $A_1,\dots,A_n\in U$, let us construct words $g_i(\alpha)$, $g_i(\alpha(0))$ being ε- close to A_i, such that the values $s_i(u)=\sigma(g_i(\alpha(u)))$, $i=1,\dots,n$, are functions of joint rank n at 0. This will prove Lemma 2.5.

Given a tangent vector $v_1\in T_0\hat{U}\setminus 0$, there exists a word g_1 (denote $s_1(u)=\sigma(g_1(\alpha(u))))$ such that $g_1(\alpha(0))$ is ε- close to A_1 and $\frac{ds_1(u)}{dv_1}\neq 0$ (*conj*- nondegeneracy and Corollary 2.12 applied to $g=A_1$). Take another vector $v_2\neq 0$ tangent to the level hypersurface of the function s_1 at 0. Again applying the corollary, one can find a word g_2 with $g_2(\alpha(0))$ being ε- close to A_2 such that the derivative along v_2 of the function $s_2: u\mapsto\sigma(g_2(\alpha(u)))$ does not vanish. Now take a vector $v_3\neq 0$ tangent to the level surface of the vector function (s_1,s_2) and construct a word g_3 similarly etc. This yields the words g_i we are looking for: by construction, the system of functions $s_i: u\mapsto\sigma(g_i(\alpha(u)))$ has rank n at 0. Lemma 2.5 is proved modulo Lemma 2.11 and Corollary 2.12. □

In the proofs of Lemma 2.11 and Corollary 2.12 we use the following notation. For any $g\in G$ and any tangent vector $v\in T_gG$ we consider its extension up to a left-invariant vector field on G and denote this field by the same symbol v. We use the following well-known derivation rule (here the derivatives (which are tangent vectors) are treated as the extended left-invariant vector fields):

for any two element families $a,b:\hat{U}\to G$ and a vector $v\in T_0\hat{U}$ one has

$$\frac{d(ab(u))}{dv}=\frac{db(u)}{dv}+Ad_{b(0)}^{-1}\frac{da(u)}{dv}. \tag{2.22}$$

More generally, for any $v\in T_0\hat{U}$, $k\in\mathbb{N}$ and any mappings $\psi_1,\dots,\psi_k:\hat{U}\to G$ one has

$$\frac{d}{dv}(\psi_1(u)\dots\psi_k(u))=\frac{d\psi_k(u)}{dv}+Ad_{\psi_k^{-1}(0)}\frac{d\psi_{k-1}(u)}{dv}+\dots+Ad_{\psi_k^{-1}(0)\dots\psi_2^{-1}(0)}\frac{d\psi_1(u)}{dv}. \tag{2.23}$$

Proof of Corollary 2.12. Without loss of generality we consider that $d\sigma(g)\neq 0$ (local nonconstance of σ). By density, it suffices to prove the corollary for a g represented by some word $w(\alpha(0))$. If the derivative $\frac{dw(\alpha(u))}{dv}$ is already transversal to $\sigma=const$, then we are done: we put $\widetilde{w}_k=w$. Suppose now that it is tangent to the level of σ. Let us modify w to make the derivative transversal.

Let $\Lambda'\subset T_{w(\alpha(0))}G$ be a line transversal to the level hypersurface of σ passing through $w(\alpha(0))$. Denote $\Lambda\subset\mathfrak{g}=T_1G$ the image of the line Λ' under the left multiplication by $w^{-1}(\alpha(0))$. Let w_k be the corresponding words from Lemma 2.11. We claim that the words $\widetilde{w}_k=ww_k$ are those we are looking for. Indeed, denote

$g=w(\alpha(0))$, $h_k=\widetilde{w}_k(\alpha(0))$. One has $h_k\to g$ by construction. Denote

$$\nu=\frac{dw(\alpha(u))}{dv},\ \nu_k=\frac{dw_k(\alpha(u))}{dv},\ \widetilde{\nu}_k=\frac{d\widetilde{w}_k(\alpha(u))}{dv}$$

and consider their left-invariant field extensions denoted by the same symbols. One has

$$\widetilde{\nu}_k = \nu_k + Ad^{-1}_{w_k(\alpha(0))}\nu = \nu_k + \nu + O(dist(w_k(\alpha(0)),1)) = \nu_k + \nu + o(\nu_k), \text{ as } k \to \infty. \quad (2.24)$$

This follows from formula (2.22) applied to $a = w(\alpha(u))$, $b = w_k(\alpha(u))$ and (2.21).

For any vector field V on G and any $h \in G$ we denote $d\sigma(V)(h)$ the value at h of the derivative of σ along V. For the proof of the corollary it suffices to show that $d\sigma(\widetilde{\nu}_k)(h_k) \neq 0$, whenever k is large enough. To do this, we show that the derivative $d\sigma(\nu_k)(h_k)$ is nonzero and asymptotically dominates the derivatives of σ along the other terms in the right-hand side of (2.24). One has $d\sigma(\nu)(g) = 0$ by definition. Hence,

$$d\sigma(\nu)(h_k) = O(dist(h_k, g)) = O(dist(w_k(\alpha(0)),1)) = o(\nu_k),$$

by (2.21). There exists a constant $c > 0$ such that $|d\sigma(\nu_k)(h_k)| > c|\nu_k|$ for all k large enough (this implies the previous asymptotic domination statement). Indeed, the vector of the field ν_k at h_k generates a line in $T_{h_k}G$ that tends to the line $\Lambda' \subset T_gG$ (which is transversal to the level hypersurface of σ by definition). This follows from the left-invariance of the field ν_k and the fact that the vector of ν_k at $w_k(\alpha(0))$ generates a line tending to Λ (Lemma 2.11). This implies the previous inequality. Now the previous domination statement implies that $d\sigma(\widetilde{\nu}_k)(h_k) \neq 0$ for large k. This proves the corollary. □

Proof of Lemma 2.11. Firstly we prove statement 1): let us show that one can always find a word sequence w_k, $w_k(\alpha(0)) \to 1$, so that the corresponding derivatives are not Lipschitz at 1:

$$dist(w_k(\alpha(0)),1) = o(|\frac{dw_k(\alpha(u))}{dv}|), \text{ as } k \to \infty.$$

If the vector field on $\Gamma =< \alpha(0) >$ from the beginning of the subsection is not single-valued, the previous statement obviously holds true: there exists a word w giving a relation in Γ, $w(\alpha(0)) = 1$, and presenting a mapping $u \mapsto w(\alpha(u))$ with nonzero derivative along v (see the beginning of the subsection); one can put $w_k = w$. Thus, everywhere below in this subsection without loss of generality we assume that the vector field is well-defined on Γ. Afterwards, using the irreducibility of the adjoint and density, we deduce that the line generated by the derivative can approach arbitrary given line in $\mathfrak{g}$. This will prove the lemma.

We prove statement 1) by contradiction. Suppose the contrary: the vector field is Lipschitz at 1, i.e., its vectors at points $w(\alpha(0))$ are $O(dist(w(\alpha(0)),1))$, as $w(\alpha(0)) \to 1$. We claim that it extends up to a vector field on G that is locally Lipschitz at each point (with uniform Lipschitz constants on compact subsets).

By density, it suffices to show that for any compact subset $K \subset G$ (say, a ball) there exists a constant $c > 0$ such that the previous vector field is c- Lipschitz on $K \cap \Gamma$. This is equivalent to say that there exists a constant $\hat{c} > 0$ such that each point of $K \cap \Gamma$ has a neighborhood in K where the vector field is $\hat{c}$- Lipschitz. Or in other terms (by compactness), for any two word sequences w_k, $\widetilde{w}_k$ such that $h_k = w_k(\alpha(0)), \widetilde{h}_k = \widetilde{w}_k(\alpha(0)) \in K$ and $dist(h_k, \widetilde{h}_k) \to 0$, as $k \to \infty$, the corresponding derivatives (more precisely, their left-invariant field extensions) differ at h_k by a quantity of order $O(dist(h_k, \widetilde{h}_k))$. Let us fix the previous word sequences w_k and $\widetilde{w}_k$ and prove the latter statement. Without loss of generality we consider that their values converge (passing to a subsequence, by compactness): $h_k, \widetilde{h}_k \to g \in G$, as $k \to \infty$. Consider the ratios $\hat{w}_k(\alpha(u)) = \widetilde{w}_k(\alpha(u))w_k^{-1}(\alpha(u))$, whose values at $u = 0$ tend to 1. By the previous Lipschitz assumption at 1, their derivatives along v tend to 0 and are

$O(dist(\widetilde{h}_k h_k^{-1}, 1)) = O(dist(h_k, \widetilde{h}_k))$. This implies that the derivatives along v of $w_k(\alpha(u))$ and $\widetilde{w}_k(\alpha(u))$ (more precisely, their left-invariant field extensions) differ by a quantity with a similar asymptotics (by applying (2.22) to $a(u) = \hat{w}_k(\alpha(u))$, $b(u) = w_k(\alpha(u))$). The Lipschitz property is proved. It implies that the field of derivatives extends up to a locally Lipschitz vector field on G by passing to limits (the density of Γ).

The flow of the latter field is well-defined (at least locally near 1, by its Lipschitz property). The vector field agrees with the multiplication. This means that for any two elements $a, b \in G$ (denote v_a, v_b, v_{ab} the corresponding field vectors at a, b and ab respectively) the vector v_{ab} is exactly the infinitesimal movement vector of the element ab, while a and b move infinitesimally along v_a and v_b respectively (or equivalently, formula (2.22) holds true for the derivatives along the flow). Indeed, this is true on the subgroup Γ (by definition), which is dense. Therefore, the flow of the previous field is given by (local) automorphisms of G. Since G is semisimple, any flow of automorphisms preserves conjugacy classes. Therefore, the mapping $u \mapsto Conj(\alpha(u))$ has zero derivative along v, - a contradiction.

Now let us prove the second statement of the lemma. Consider all the sequences w_k, $w_k(\alpha(0)) \to 1$, that satisfy (2.21). Consider the lines in $T_{w_k(\alpha(0))}G$ generated by the derivatives along v of $w_k(\alpha(u))$ and all the limits of these lines (along subsequences). The limit lines lie in $\mathfrak{g}$. Denote $\hat{L} \subset \mathfrak{g}$ their union. It is nonempty by the compactness of projective space and closed by definition.

Below we show that $\hat{L} = \mathfrak{g}$: this will prove that statement 2) of Lemma 2.11 holds true for appropriate sequence w_k.

Claim. *The set $\hat{L}$ is Ad_G- invariant.*

Proof It suffices to prove the $Ad_{w(\alpha(0))}$- invariance of $\hat{L}$ for any word w (the density of Γ). Let us fix a w, a line $\Lambda \subset \hat{L}$ and show that

$$\Lambda_w = Ad_{w(\alpha(0))}\Lambda \subset \hat{L}. \tag{2.25}$$

To do this, fix a word sequence w_k satisfying the statements of Lemma 2.11 for the line Λ (it exists by definition). The word sequence $\widetilde{w}_k = w w_k w^{-1}$ satisfies its statements for the line Λ_w. Indeed, the left-invariant field extensions of the derivatives $\nu_k = \frac{dw_k(\alpha(u))}{dv}$ and $\widetilde{\nu}_k = \frac{d\widetilde{w}_k(\alpha(u))}{dv}$ satisfy the asymptotic formula

$$\widetilde{\nu}_k = Ad_{w(\alpha(0))}\nu_k + O(dist(w_k(\alpha(0)), 1)) = Ad_{w(\alpha(0))}\nu_k + o(\nu_k), \tag{2.26}$$

by (2.22) (applied to the equal products $\widetilde{w}_k w = w w_k$) and (2.21). This implies that the words $\widetilde{w}_k$ satisfy (2.21):

$$dist(\widetilde{w}_k(\alpha(0)), 1) = O(dist(w_k(\alpha(0)), 1)) = o(\nu_k) = o(\widetilde{\nu}_k).$$

The line in $T_{w_k(\alpha(0))}G$ generated by the left-invariant vector field ν_k tends to Λ, as $k \to \infty$ (Lemma 2.11). (This is equivalent to the similar statement for the corresponding line in T_1G, since $w_k(\alpha(0)) \to 1$ and by left-invariance.) This together with (2.26) implies the similar statements for the vector field $\widetilde{\nu}_k$ and the line Λ_w and proves (2.25) and the Claim. □

Fix a line $\Lambda \subset \hat{L}$ and a vector $\nu \in \Lambda \setminus 0$. The adjoint Ad_G is irreducible. Hence, there exists a collection of $n-1$ elements $g_1, \dots, g_{n-1} \in G$ such that the vectors $\nu, Ad_{g_1}\nu, \dots, Ad_{g_{n-1}}\nu \subset \hat{L}$

(see the Claim) generate $\mathfrak{g}$ as a linear space (we put $g_0 = 1$). By density, the latter elements g_i, $i \geq 1$, can be chosen to be the values

$$g_i = \omega_i(\alpha(0)) \tag{2.27}$$

of appropriate words ω_i. For any collection $r = (r_0, \dots, r_{n-1})$ of integer numbers that do not vanish simultaneously we show that

$$\text{the line } \Lambda(r) = \mathbb{R}(\sum_{i\geq 0} r_i Ad_{g_i}\nu) \text{ is contained in } \hat{L}. \tag{2.28}$$

The union of all the lines $\Lambda(r)$ is dense in $\mathfrak{g}$, since $Ad_{g_i}\nu$ is a basis. This together with the closeness of $\hat{L}$ implies that $\hat{L} = \mathfrak{g}$.

For the proof of (2.28) we fix a sequence of words w_k satisfying the statements of Lemma 2.11 for the line Λ (it exists by definition and since $\Lambda \subset \hat{L}$). Consider the following auxiliary word sequence:

$$\widetilde{w}_k = w_k^{r_0}(\omega_1 w_k \omega_1^{-1})^{r_1} \dots (\omega_{n-1} w_k \omega_{n-1}^{-1})^{r_{n-1}}, \text{ where } \omega_i \text{ are the same, as in (2.27).}$$

We claim that the derivatives $\widetilde{\nu}_k = \frac{d\widetilde{w}_k(\alpha(u))}{dv}$ satisfy (2.21) and generate lines tending to $\Lambda(r)$ (hence, $\Lambda(r) \subset \hat{L}$). Indeed, the left-invariant field extensions of the derivatives $\nu_k = \frac{dw_k(\alpha(u))}{dv}$ and $\widetilde{\nu}_k$ satisfy the asymptotic formula

$$\widetilde{\nu}_k = (r_0 + \sum_{i\geq 1} r_i Ad_{\omega_i(\alpha(0))})\nu_k + o(\nu_k), \tag{2.29}$$

which follows from (2.26) (with $\widetilde{w}_k$ replaced by $\chi_j = \omega_j w_k \omega_j^{-1}$), (2.23) (applied to the above product $\widetilde{w}_k$ of $w_k^{r_0}$ and $\chi_j^{r_j}$) and (2.21). Formula (2.29) implies the previous statements on the derivatives $\widetilde{\nu}_k$ and $\Lambda(r)$, as in the proof of the previous Claim. This proves (2.28). This together with the previous discussion proves that $\hat{L} = \mathfrak{g}$. Lemma 2.11 is proved. The proof of Lemma 2.5 is complete. □

3 Case of semisimple Lie groups with irreducible adjoint and without proximal elements

3.1 The plan of the proof of Theorem 1.33

In the case mentioned in the title of the section the proof (given below) of Theorem 1.33 is essentially the same, as before, but it becomes slightly more technical.

Everywhere below in this section, whenever the contrary is not specified, we consider that G is a semisimple Lie group with irreducible adjoint and no proximal elements. Let $\alpha(u) = (a_1(u), \dots, a_M(u))$ be a *conj*- nondegenerate family of M- ples of its elements. As in Section 2, we consider that $u \in \mathbb{R}^n$, $n = dimG$. We construct appropriate sequence of words w_r and a sequence $u_r \in \mathbb{R}^n$ such that

$$w_r(\alpha(u_r)) = 1, \; u_r \to 0, \text{ as } r \to \infty, \tag{3.1}$$

and the relations $w_r(\alpha(u)) = 1$ do not hold identically in a neighborhood of 0. This will prove Theorem 1.33.

We construct appropriate words $g_1, \dots, g_n$, h, w, a collection

$$l = (l_1, \dots, l_n) \in \mathbb{Z}^n, \text{ a sequence of numbers } k_r \in \mathbb{N}, \ k_r \to \infty, \text{ as } r \to \infty,$$

a collection of sequences

$$m_{jr} \in \mathbb{N}, \ j = 1, \dots, n, \ r \in \mathbb{N}, \text{ and put}$$

$$\omega_r = w_{1,k_r+l_1}^{m_{1r}} \dots w_{n,k_r+l_n}^{m_{nr}}, \ w_r = w^{-1}\omega_r, \tag{3.2}$$

where w_{j,k_r+l_j} are the iterated commutators given by the recurrent formula (2.2). We consider the rescaled parameter

$$\widetilde{u} = k_r u \text{ and show that}$$

$$\omega_r(\alpha(k_r^{-1}\widetilde{u})) \to \Psi(\widetilde{u}), \ \Psi : \mathbb{R}^n \to G \text{ is a local diffeomorphism at } 0, \tag{3.3}$$

the latter convergence is uniform with derivatives on compact sets in $\mathbb{R}^n$. This implies Theorem 1.33 analogously to the discussion at the end of Subsection 2.1. The implication is proved at the end of the present subsection.

In the proof of Theorem 1.33 we use Proposition 3.8 stated below and proved in 3.3. It describes the asymptotic behavior of iterated commutators

$$\phi_g^k(h) = [g \dots [g, h] \dots],$$

as $k \to \infty$. This is an analogue of Proposition 2.2 from 2.1. In the case under consideration the unity component of G contains no 1- proximal elements (for which Proposition 2.2 was formulated). We introduce so-called *$\mathbb{C}$-1-proximal elements* (see the next definition). We show that their set contains an open dense subset in the unity component (Proposition 3.1 and its Corollary 3.4, both stated below). Proposition 3.1 is proved in 3.2. We state and prove Proposition 3.8 for the $\mathbb{C}$-1-proximal elements g such that the derivative $\phi_g'(1)$ is contracting. To do this, we show (Proposition 3.5 below) that for each $\mathbb{C}$-1-proximal element $g \in G$ there exists a unique $\phi_g'(1)$- invariant plane $L(g) \subset \mathfrak{g}$ equipped with a natural $\phi_g'(1)$-invariant complex structure such that the restriction $\phi_g'(1) : L(g) \to L(g)$ is multiplication by a complex eigenvalue $s(g)$ of the operator $\phi_g'(1) : \mathfrak{g} \to \mathfrak{g}$ with maximal modulus.

The words g_j will be chosen at the end of the subsection, in particular, so that each element $g = g_j(\alpha(0))$ be $\mathbb{C}$-1-proximal and $|s(g)| < 1$. For any collection of words g_j satisfying the latter statements and any given $\varepsilon > 0$, Proposition 3.9 and Corollary 3.10 (both stated below) provide sequences $k_r, m_{jr} \to \infty$ such that for any word h with $h(\alpha(0))$ close enough to the unity and any collection $l = (l_1, \dots, l_n) \in \mathbb{Z}^n$ the corresponding sequence of G- valued functions $\omega_r(\alpha(k_r^{-1}\widetilde{u}))$, see (3.2), converges to some mapping $\Psi : \mathbb{R}^n \to G$ uniformly with derivatives on compact sets in $\mathbb{R}^n$. The limit mapping Ψ is given explicitly by formula (3.11) below, which depends only on the words g_j, h, the collection $l \in \mathbb{Z}^n$ and ε. We show (Lemma 3.11 below) that one can adjust g_j, h and l so that Ψ be a local diffeomorphism at 0, whenever ε is small enough. This is the main technical part of the proof of Theorem 1.33. The proof of Lemma 3.11 given in 3.4 uses the Main Technical Lemma from Subsection 2.1.

At the end of the subsection we deduce Theorem 1.33 from the technical statements listed above. Afterwards we formulate Theorem 3.12, which summarizes the technical results of Sections 2 and 3. It will be used in the proof of Theorem 1.1 in the general case (Section 4) and in the proof of Theorem 1.29 (Section 6).

Proposition 3.1 *Let G be a connected semisimple Lie group. There exists a nonempty subset $U \subset G$ such that the subset $Ad_U \subset Ad_G \subset End(\mathfrak{g})$ is Zariski open in Ad_G and the adjoint of each $g \in U$ satisfies the following statements:*

1) the number of its nonunit complex eigenvalues is maximal and nonempty, and all they are simple;

2) if there is a pair of distinct eigenvalues $\Lambda_1, \Lambda_2 \neq 1$ with $|\Lambda_1 - 1| = |\Lambda_2 - 1|$, then $\Lambda_1 = \overline{\Lambda}_2$.

The proposition is proved in 3.2.

Definition 3.2 An element g of a Lie group is called $\mathbb{C}$-*1-proximal*, if the operator $Ad_g - Id$ has a pair of simple nonreal complex-conjugated eigenvalues that are the unique eigenvalues with maximal modulus.

Proposition 3.3 *Any element of a semisimple Lie group whose adjoint satisfies the previous statements 1) and 2) is either 1- proximal (see Definition 1.56) or $\mathbb{C}$-1-proximal.*

Proof Let Ad_g satisfy 1) and 2), λ be its eigenvalue for which the modulus $|\lambda - 1|$ is the maximal possible. Then $\lambda - 1 \neq 0$ and λ is a simple eigenvalue (statement 1)). For any eigenvalue $\lambda' \neq \lambda, \bar{\lambda}$ one has $|\lambda - 1| > |\lambda' - 1|$ (statement 2)). Therefore, g is 1- proximal, if $\lambda \in \mathbb{R}$ and $\mathbb{C}$-1-proximal otherwise. Proposition 3.3 is proved. □

Corollary 3.4 *Let G be a semisimple Lie group without proximal elements. The set of $\mathbb{C}$- 1-proximal elements in G is open and contains a dense subset $U \subset G_0$ of its unity component G_0.*

Proof The openness of the set of $\mathbb{C}$-1-proximal elements follows from definition. The subset $U \subset G_0$ from Proposition 3.1 is open and dense (since Ad_U is Zariski dense in Ad_{G_0}, by Proposition 3.1). The set U consists of $\mathbb{C}$-1-proximal elements (Proposition 3.3 and absense of 1- proximal elements in G_0). Indeed, otherwise, a 1- proximal element of G_0 would be proximal (Corollary 1.59), - a contradiction to the conditions of Corollary 3.4. This proves Corollary 3.4. □

We use the following properties of the adjoint of a $\mathbb{C}$-1-proximal element.

Proposition 3.5 *Let $A : \mathbb{R}^n \to \mathbb{R}^n$ be a linear operator with a pair of simple complex-conjugated eigenvalues $s, \bar{s} \notin \mathbb{R}$. There exists a unique A- invariant plane $L \subset \mathbb{R}^n$ whose complexification is the sum of the complex eigenlines corresponding to the eigenvalues s and $\bar{s}$. The plane L carries an A- invariant linear complex structure (i.e., a structure of complex line compatible with its real linear structure), unique up to complex conjugation. The restriction $A : L \to L$ acts by multiplication by either s or $\bar{s}$ in the latter complex structure (dependently on double choice of the eigenvalue).*

Proof By basic linear algebra, the previous plane L exists, unique and there exists a $\mathbb{R}$-linear nondegenerate operator $H : L \to \mathbb{C}$ such that $HAH^{-1}(z) = sz$. The H- pullback of the standard complex structure on $\mathbb{C}$ (or of its conjugate) is an A- invariant complex structure on L such that the restriction $A : L \to L$ acts by multiplication by s (respectively, $\bar{s}$). These are the only A- invariant linear complex structures on L. Or equivalently, the standard complex

structure on $\mathbb{C}$ is the unique linear complex structure (up to complex conjugation) invariant under the multiplication by a number $s \in \mathbb{C} \setminus \mathbb{R}$. Indeed, each linear complex structure on a plane defines an ellipse centered at 0 (up to homothety): the latter ellipse is an orbit of a vector under the multiplication by the complex numbers with unit modulus. Vice versa, an ellipse determines a linear complex structure uniquely up to complex conjugation. The only ellipse in $\mathbb{C}$ sent to a homothetic one by multiplication by a $s \in \mathbb{C} \setminus \mathbb{R}$ is a circle. This proves the previous uniqueness statement and Proposition 3.5. $\square$

Definition 3.6 Let G be a Lie group, $g \in G$ be a $\mathbb{C}$-1-proximal element. Let $s(g)$ be an eigenvalue of $Ad_g - Id$ with the maximal modulus. Let $L(g) \subset \mathfrak{g}$ be the $Ad_g - Id$- (and hence, Ad_g-) invariant plane corresponding to the eigenvalues $s(g)$, $\overline{s(g)}$ (see Proposition 3.5). The corresponding $Ad_g - Id$- invariant complex structure on $L(g)$, in which $Ad_g - Id : L(g) \to L(g)$ acts by multiplication by $s(g)$, will be called the $s(g)$- *complex structure.*

Proposition 3.7 *Let G be a Lie group, $V \subset G$ be a connected component of the subset of the $\mathbb{C}$-1-proximal elements (which is open by definition). The values $s(g)$, $\overline{s(g)}$ from Definition 3.6 yield two real-analytic complex-conjugated functions $s, \bar{s} : V \to \mathbb{C} = \mathbb{R}^2$.*

Proof The local real analyticity of the previous values follows from the simplicity of the eigenvalues $s(g)$, $\overline{s(g)}$. The global real analyticity (say, of $s(g)$) follows from the fact that its analytic extension along any closed loop in V does not change the analytic branch. Indeed, the result of analytic extension of $s(g)$ remains an eigenvalue of $Ad_g - Id$ with the maximal modulus, by definition and the previous local analyticity statement. Therefore, given a $g_0 \in V$ and a loop $\gamma \subset V$ based at g_0, the result of the analytic extension of $s(g)$ along γ is either $s(g_0)$, or $\overline{s(g_0)}$. In the latter case there exists a $g' \in \gamma$ where $s(g') \in \mathbb{R}$, by continuity. It follows from definition and the local analyticity that $s(g')$ is a double eigenvalue of $Ad_{g'} - Id$ with maximal modulus, - a contradiction to the $\mathbb{C}$-1- proximality. Proposition 3.7 is proved. $\square$

In what follows, everywhere below in this Section, we fix a real-analytic branch of the eigenvalue function $s(g)$ from Proposition 3.7, defined on the open set of all the $\mathbb{C}$-1-proximal elements. The corresponding family of planes $L(g) \subset \mathfrak{g}$ and the $s(g)$- complex structures on them (see the previous Definition) also depend analytically on g. We define the multiplication of vectors in $L(g)$ by complex numbers in the sense of the $s(g)$- complex structure. Denote

$$\Pi_{\mathbb{C},1} = \{\mathbb{C} - 1 - \text{proximal elements } g \in G \text{ with } |s(g)| < 1\} \tag{3.4}$$

This is a nonempty open subset in G, by Corollary 3.4.

Proposition 3.8 *Let G be a Lie group such that $\Pi_{\mathbb{C},1} \neq \emptyset$, $s(g)$, $L(g)$ and the complex structures on the planes $L(g)$ be as above. There exists an open subset*

$$\Pi'_{\mathbb{C},1} \subset \Pi_{\mathbb{C},1} \times G, \ \Pi_{\mathbb{C},1} \times 1 \subset \Pi'_{\mathbb{C},1}, \tag{3.5}$$

and a $\mathfrak{g}$- valued vector function $v_g(h)$ analytic in $(g,h) \in \Pi'_{\mathbb{C},1}$ (denote $dv_g : \mathfrak{g} \to \mathfrak{g}$ its differential in h at $h = 1$) such that

$$v_g(1) = 0, \ v_g(h) \in L(g) \textit{ for any } (g,h) \in \Pi'_{\mathbb{C},1}, \ dv_g|_{L(g)} = Id : L(g) \to L(g), \tag{3.6}$$

$$\phi_g^k(h) = \exp(s^k(g)v_g(h) + o(|s^k(g)|)), \text{ as } k \to \infty, \tag{3.7}$$

the latter "o" is uniform with derivatives on compact subsets in $\Pi'_{\mathbb{C},1}$.

Proposition 3.8 is proved in 3.3.

Given a collection of words g_j, $j = 1, \dots, n$, with $g_j(\alpha(0)) \in \Pi_{\mathbb{C},1}$, we denote

$$\zeta_j = \arg s(g_j(\alpha(0))).$$

Proposition 3.9 *For any real vector* $\zeta = (\zeta_1, \dots, \zeta_n) \in \mathbb{R}^n$ *there exists a sequence of numbers* $k_r \in \mathbb{N}$, $k_r \to \infty$, *as* $r \to \infty$, *such that*

$$k_r\zeta_j \to 0(mod 2\pi), \text{ as } r \to \infty, \text{ for any } j = 1, \dots, n. \tag{3.8}$$

Proof Consider ζ as an element of the torus $\mathbb{T}^n = \mathbb{R}^n/2\pi\mathbb{Z}^n$. The subgroup $< \zeta > \subset \mathbb{T}^n$ either is discrete, or accumulates to 0. In both cases there exists a sequence of numbers $k_r \in \mathbb{N}$, $k_r \to \infty$, such that $k_r\zeta \to 0$ in $\mathbb{T}^n$ (the latter statement is equivalent to (3.8)). In the second case this follows from definition. In the first case the group $< \zeta >$ is finite cyclic by compactness. Denote m its order, $k_r = rm$. Then $k_r\zeta = 0$ in $\mathbb{T}^n$ for all $r \in \mathbb{N}$. This proves Proposition 3.9. □

Corollary 3.10 *Let* G, n, M, $\alpha(u)$ *be as at the beginning of the Subsection,* $\Pi_{\mathbb{C},1}$ *be as in (3.4),* $\Pi'_{\mathbb{C},1}$ *be as in (3.5). Let* $g_1, \dots, g_n, h$ *be words in* M *elements such that*

$$(g_j(\alpha(0)), h(\alpha(0))) \in \Pi'_{\mathbb{C},1} \text{ for any } j = 1, \dots, n. \tag{3.9}$$

Let $k_r \in \mathbb{N}$, $k_r \to \infty$, *be a sequence satisfying (3.8) with* $\zeta_j = \arg s(g_j(\alpha(0)))$. *Let* $\varepsilon > 0$, *put*

$$m_{jr} = [\varepsilon|s|^{-k_r}(g_j(\alpha(0)))], \ s_j(u) = s(g_j(\alpha(u))), \ \nu_j = v_{g_j(\alpha(0))}(h(\alpha(0))) \in L(g_j(\alpha(0))), \tag{3.10}$$

see (3.6). Let $l = (l_1, \dots, l_n) \in \mathbb{Z}^n$ *be an arbitrary collection of* n *integers,* ω_r *be the corresponding product (3.2) of iterated commutator powers. Then*

$$\omega_r(\alpha(k_r^{-1}\widetilde{u})) \to \Psi(\widetilde{u}) = \exp(\varepsilon s_1^{l_1}(0)e^{(d\ln s_1(0))\widetilde{u}}\nu_1)\dots\exp(\varepsilon s_n^{l_n}(0)e^{(d\ln s_n(0))\widetilde{u}}\nu_n), \tag{3.11}$$

as $r \to \infty$, *uniformly with derivatives on compact sets in* $\mathbb{R}^n$. *(The multiplication of the vectors* $\nu_j \in L(g_j(\alpha(0)))$ *by complex numbers is defined in terms of the* $s(g_j(\alpha(0)))$*- complex structures on* $L(g_j(\alpha(0)))$*.)*

Proof One has (as $r \to \infty$)

$$w_{j,k_r+l_j}(\alpha(k_r^{-1}\widetilde{u})) = \exp(s_j^{k_r+l_j}(k_r^{-1}\widetilde{u})\widetilde{\nu}_j(\widetilde{u}) + o(|s_j^{k_r+l_j}(k_r^{-1}\widetilde{u})|)), \text{ where} \tag{3.12}$$

$$\widetilde{\nu}_j(\widetilde{u}) = v_{g_j(\alpha(k_r^{-1}\widetilde{u}))}(h(\alpha(k_r^{-1}\widetilde{u}))) \to \nu_j,$$

by definition and (3.7),

$$s_j^{k_r+l_j}(k_r^{-1}\widetilde{u}) = s_j^{k_r+l_j}(0)e^{(d\ln s_j(0))\widetilde{u}}(1 + o(1)), \text{ as in (2.16)},$$

$m_{jr}s_j^{k_r}(0) \to \varepsilon$ by (3.8) and (3.10). Hence, $w_{j,k_r+l_j}^{m_{jr}}(\alpha(k_r^{-1}\widetilde{u})) \to \exp(\varepsilon s_j^{l_j}(0)e^{(d\ln s_j(0))\widetilde{u}}\nu_j)$, as $r \to \infty$, by (3.12) and the latter asymptotics. This implies (3.11). □

Lemma 3.11 *Let G, n, M, $\alpha(u)$ be as at the beginning of the subsection. There exist words $g_1, \dots, g_n, h$ satisfying (3.9) and a $l = (l_1, \dots, l_n) \in \mathbb{Z}^n$ such that for any $\varepsilon > 0$ small enough the corresponding mapping $\Psi(\widetilde{u})$ from (3.11) be a local diffeomorphism at 0.*

Lemma 3.11 is proved in 3.4.

Proof of Theorem 1.33 modulo Propositions 3.1, 3.8 and Lemma 3.11. Let $g_1, \dots, g_n$, h, l, ε be as in Lemma 3.11, $s_j(u) = s(g_j(\alpha(u)))$, $\zeta_j = \arg s_j(0)$. Let $k_r \to \infty$ be a natural sequence satisfying (3.8). Let m_{jr} be the numbers from (3.10). Let ω_r be the corresponding iterated commutator power product (3.2), Ψ be the corresponding mapping from (3.11). Let $\delta > 0$ be such that

$$\Psi : \overline{D}_\delta \to \Psi(\overline{D}_\delta) \subset G \text{ be a diffeomorphism.}$$

It exists by Lemma 3.11. Fix an arbitrary word w in M elements such that

$$w(\alpha(0)) \in \Psi(D_\delta). \text{ Put } w_r = w^{-1}\omega_r.$$

For any r large enough the image $w_r(\alpha(k_r^{-1}D_\delta))$ contains 1. This follows from the convergence

$$w_r(\alpha(k_r^{-1}\widetilde{u})) \to \psi(\widetilde{u}) = w^{-1}(\alpha(0))\Psi(\widetilde{u}) \tag{3.13}$$

(which takes place by definition and (3.11)) and the fact that

$$\psi : \overline{D}_\delta \to \psi(\overline{D}_\delta) \subset G \text{ is a diffeomorphism, and } 1 \in \psi(D_\delta),$$

as at the end of Subsection 2.1. Therefore, for any r large enough there exists a parameter value

$$\widetilde{u}_r \in D_\delta \subset \mathbb{R}^n, \text{ put } u_r = k_r^{-1}\widetilde{u}_r, \text{ such that } w_r(\alpha(u_r)) = 1.$$

The sequence u_r satisfies (3.1). The relations $w_r(\alpha(u)) = 1$ do not hold identically in u for any r large enough, as at the end of 2.1. This proves Theorem 1.33 modulo Propositions 3.1, 3.8 and Lemma 3.11. □

The next theorem summarizes the technical results of Sections 2 and 3. It will be used in the proof of Theorem 1.1 in the general case and in the proof of Theorem 1.29.

Theorem 3.12 *Let G be a semisimple Lie group with irreducible adjoint, $n = dimG$. Let $M \in \mathbb{N}$, $\alpha(u) = (a_1(u), \dots, a_M(u)) \in G^M$ be a conj- nondegenerate family depending on the parameter $u \in \mathbb{R}^n$. There exist a sequence of words $w_r(a_1, \dots, a_M)$, a sequence of numbers $k_r \in \mathbb{N}$, $k_r \to \infty$, as $r \to \infty$ ($k_r = r$, if G has proximal elements), a smooth (real-analytic, depending on the regularity of $\alpha(u)$) mapping $\psi : \mathbb{R}^n \to G$ and a $\delta > 0$ such that*

$$w_r(\alpha(k_r^{-1}\widetilde{u})) \to \psi(\widetilde{u}) \text{ uniformly with derivatives on compact subsets, as } r \to \infty, \tag{3.14}$$

$$\psi : \overline{D}_\delta \to \psi(\overline{D}_\delta) \subset G \text{ is a diffeomorphism and } 1 \in \psi(D_\delta). \tag{3.15}$$

Theorem 3.12 follows from statement (2.17) (if G has proximal elements) or (3.13) otherwise.

Definition 3.13 In the conditions of Theorem 3.12 the tuple $(\{w_r\}, \{k_r\}, \psi, \delta)$ is called the *converging tuple* associated to the given *conj*- nondegenerate family $\alpha(u)$.

3.2 Complex eigenvalues of adjoint and $\mathbb{C}$-1-proximal elements

Here we prove Proposition 3.1.

Denote

$$r = \text{ the rank of } G, \text{ i.e., the minimal multiplicity of zero eigenvalue of } ad_x, \ x \in \mathfrak{g}.$$

The characteristic polynomial $\chi_{\mathrm{Ad}_g}(\lambda)$ of the adjoint Ad_g, $g \in G$, has the type

$$\chi_{Ad_g}(\lambda) = (1-\lambda)^r Q_g(\lambda), \ Q_g(\lambda) = \lambda^{n-r} + 1 + \sum_{j=1}^{n-r-1} q_j(g)\lambda^j,$$

where $q_j(g)$ are polynomials in the matrix coefficients of the operator Ad_g (in some fixed basis of $\mathfrak{g}$). The free term of the polynomial Q_g is unit, since $\chi_{Ad_g}(0) = \det Ad_g = 1$, by semisimplicity, see, e.g., Proposition 1.58. Its higher coefficient equals $(-1)^{n-r} = 1$, by definition and since the number $n-r$ is even. This follows from the same proposition. Denote

$$\lambda_1(g), \dots, \lambda_{n-r}(g) \text{ the complex roots of the polynomial } Q_g(\lambda),$$

which are multivalued functions in g. Consider the following auxiliary subsets $\Lambda_0, \dots, \Lambda_4 \subset G$:

$$\Lambda_0 = \{g \in G \mid \lambda_j(g) = 1 \text{ for some } j\},$$

$$\Lambda_1 = \{g \in G \mid \lambda_{j_1}(g) = \lambda_{j_2}(g) \text{ for some } j_1 \neq j_2\},$$
$$\Lambda_2 = \{g \in G \mid \lambda_{j_1}(g) - 1 = 1 - \lambda_{j_2}(g) \text{ for some } j_1 \neq j_2\},$$
$$\Lambda_3 = \{g \in G \mid (\lambda_{j_1}(g) - 1)^2 = (\lambda_{j_2}(g) - 1)(\lambda_{j_3}(g) - 1)$$
$$\text{for some distinct indices } j_1, j_2, j_3\},$$
$$\Lambda_4 = \{g \in G \mid (\lambda_{j_1}(g) - 1)(\lambda_{j_2}(g) - 1) = (\lambda_{j_3}(g) - 1)(\lambda_{j_4}(g) - 1)$$
$$\text{for some distinct indices } j_1, \dots, j_4\}.$$

Put

$$U = G \setminus \cup_{i=0}^4 \Lambda_i.$$

We introduce on G the Zariski topology: the pullback of that on $End(\mathfrak{g})$ under the adjoint representation.

The set U is Zariski open, since the sets Λ_i are Zariski closed by definition. (A priori, a Zariski open subset may be empty.) Each $g \in U$ satisfies statements 1) and 2) of Proposition 3.1. Indeed, if a $g \in G$ does not satisfy statement 1), then $g \in \Lambda_0 \cup \Lambda_1 \subset G \setminus U$ by definition. If a $g \in G$ does not satisfy statement 2), then there exist eigenvalues ν_1, ν_2 of Ad_g such that

$$|\nu_1 - 1| = |\nu_2 - 1| \text{ and } \nu_1 \neq \nu_2, \bar{\nu}_2, \tag{3.16}$$

which implies that $\nu_1, \nu_2 \neq 1$. Hence,

$$\nu_1 = \lambda_{j_1}(g), \ \nu_2 = \lambda_{j_2}(g), \ j_1 \neq j_2. \text{ We consider that } j_1 = 1, \ j_2 = 2, \tag{3.17}$$

without loss of generality, choosing appropriate numeration of the λ_j' s. We claim that in this case

$$g \in \Lambda_2 \cup \Lambda_3 \cup \Lambda_4 \subset G \setminus U.$$

Case $\nu_1, \nu_2 \in \mathbb{R}$. Then by (3.17), $\lambda_1(g), \lambda_2(g) \in \mathbb{R}$, $\lambda_1(g) \neq \lambda_2(g)$. Hence,

$$\lambda_1(g) - 1 = 1 - \lambda_2(g), \text{ by (3.16), thus, } g \in \Lambda_2.$$

Case $\nu_1 \in \mathbb{R}$, $\nu_2 \notin \mathbb{R}$. Denote $\lambda_3(g) = \overline{\lambda_2(g)}$, which is also a root of the real polynomial $Q_g(\lambda)$. Then $\lambda_1(g) \in \mathbb{R}$ by (3.17), and thus, by (3.16),

$$(\lambda_1(g) - 1)^2 = (\lambda_2(g) - 1)(\lambda_3(g) - 1), \text{ hence } g \in \Lambda_3.$$

Case $\nu_1, \nu_2 \notin \mathbb{R}$. Denote

$$\lambda_3(g) = \overline{\lambda_1(g)}, \ \lambda_4(g) = \overline{\lambda_2(g)}. \text{ By (3.16), (3.17),}$$

$$(\lambda_1(g) - 1)(\lambda_3(g) - 1) = (\lambda_2(g) - 1)(\lambda_4(g) - 1), \text{hence } g \in \Lambda_4.$$

This proves that any element $g \in G$ that does not satisfy some of the statements 1), 2) of Proposition 3.1 does not belong to U.

Thus, the set U is Zariski open and consists of elements satisfying statements 1) and 2). Now for the proof of Proposition 3.1 it suffices to show that the set U is nonempty. Let $\mathfrak{h} \subset \mathfrak{g}$ be an arbitrary Cartan subalgebra (see 1.9), $T = \exp \mathfrak{h} \subset G$ be the corresponding maximal torus. By definition, for any

$$g = \exp v, \ v \in \mathfrak{h}, \text{ one has } \lambda_j(g) = \exp \alpha_j(v), \ \Delta = \{\alpha_1, \dots, \alpha_{n-r}\} \text{ is the root system of } \mathfrak{h}_{\mathbb{C}}. \tag{3.18}$$

The roots α_j define distinct nonzero complex-valued linear functionals $\alpha_j : \mathfrak{h} \to \mathbb{C}$. We prove (by contradiction) a stronger statement:

$$U \cap T \neq \emptyset.$$

Suppose the contrary: $U \cap T = \emptyset$. This means that the sets $\Lambda_0, \dots, \Lambda_4$ (which are Zariski closed) cover T. One of the Λ_i' s contains T, since T is a regular connected analytic variety.

Case $T \subset \Lambda_0$. Then $\alpha_j \equiv 0$ on $\mathfrak{h}$ for some root α_j, by definition and (3.18), hence $\alpha_j = 0 \in \mathfrak{h}^*_{\mathbb{C}}$, - a contradiction.

Case $T \subset \Lambda_1$. Then $\alpha_{j_1} \equiv \alpha_{j_2}$ on $\mathfrak{h}$ for some distinct roots α_{j_1} and α_{j_2}, - a contradiction.

Case $T \subset \Lambda_2$. Then there is a pair of distinct roots, we numerate them as α_1, α_2, such that

$$e^{\alpha_1(v)} - 1 = 1 - e^{\alpha_2(v)} \text{ for any } v \in \mathfrak{h}.$$

Passing to the limit, as $v \to 0$, one gets

$$\alpha_1 \equiv -\alpha_2, \text{ thus,}$$

$$e^{\alpha_1(v)} - 1 \equiv 1 - e^{-\alpha_1(v)} \equiv e^{-\alpha_1(v)}(e^{\alpha_1(v)} - 1), \text{ and hence, } \alpha_1 \equiv 0,$$

- a contradiction.

Case $T \subset \Lambda_3$. Then there exists a triple of distinct roots, we numerate them as α_1, α_2, α_3, such that

$$(e^{\alpha_1(v)} - 1)^2 = (e^{\alpha_2(v)} - 1)(e^{\alpha_3(v)} - 1) \text{ for any } v \in \mathfrak{h}.$$

Passing to the limit, as $v \to 0$, we get

$$\alpha_1^2(v) \equiv \alpha_2(v)\alpha_3(v) \text{ for any } v \in \mathfrak{h}, \text{ and hence, for any } v \in \mathfrak{h}_{\mathbb{C}}$$

by analyticity. Therefore, the zero hyperplanes in $\mathfrak{h}_{\mathbb{C}}$ of α_1, α_2, α_3 coincide, thus, the latter roots are proportional and $\alpha_2, \alpha_3 = \pm\alpha_1$ (statement g) from Subsection 1.9). Hence, the roots $\alpha_1, \alpha_2, \alpha_3$ are not distinct, - a contradiction.

Case $T \subset \Lambda_4$. Then there exist 4 distinct roots, we numerate them as $\alpha_1, \dots, \alpha_4$, such that

$$(e^{\alpha_1(v)} - 1)(e^{\alpha_2(v)} - 1) \equiv (e^{\alpha_3(v)} - 1)(e^{\alpha_4(v)} - 1).$$

Passing to the limit, as $v \to 0$, we get (as above)

$$\alpha_1(v)\alpha_2(v) \equiv \alpha_3(v)\alpha_4(v).$$

The latter equality holds true on $\mathfrak{h}$, and hence, on $\mathfrak{h}_{\mathbb{C}}$. Therefore, the zero hyperplane in $\mathfrak{h}_{\mathbb{C}}$ of one of the roots α_3, α_4 (say, α_3) coincides with that of α_1; the zero hyperplane of α_4 coincides with that of α_2. This together with statement g) from 1.9 implies that

$$\alpha_3 = -\alpha_1, \ \alpha_4 = -\alpha_2, \text{ and hence,} \tag{3.19}$$

$$(e^{\alpha_1(v)} - 1)(e^{\alpha_2(v)} - 1) \equiv (e^{-\alpha_1(v)} - 1)(e^{-\alpha_2(v)} - 1) \equiv e^{-(\alpha_1(v)+\alpha_2(v))}(e^{\alpha_1(v)} - 1)(e^{\alpha_2(v)} - 1).$$

Therefore, $\alpha_1 \equiv -\alpha_2$. This together with (3.19) implies that $\alpha_1 = \alpha_4$, and the roots $\alpha_1, \dots, \alpha_4$ are not distinct, - a contradiction.

Thus, the intersection $U \cap T$ is nonempty. Hence, U is a nonempty Zariski open set. This proves Proposition 3.1.

3.3 Dynamics of commutator. Proof of Proposition 3.8

As it is shown below, Proposition 3.8 is implied by the following well-known fact in local dynamics.

Proposition 3.14 *Let $\phi : (\mathbb{R}^n, 0) \to (\mathbb{R}^n, 0)$ be a germ of analytic mapping, $\phi(0) = 0$. Let the differential $d\phi(0)$ have exactly two (with multiplicities) eigenvalues with maximal modulus:*

$$s, \bar{s} \in \mathbb{C} \setminus \mathbb{R}, \text{ and } 0 < |s| < 1. \text{ Let } L \subset T_0\mathbb{R}^n$$

be the invariant plane corresponding to the conjugated eigenvalue pair s, $\bar{s}$. There exist local real-analytic coordinates (z, y) on a neighborhood of zero in $\mathbb{R}^n$, $z \in \mathbb{C}$, $y \in \mathbb{R}^{n-2}$, $z(0), y(0) = 0$, such that

$$\phi(z, y) = (sz, Q(z, y)), \ Q(z, y) = Ay + O(|z|^2 + |y|^2), \tag{3.20}$$

$$A : \mathbb{R}^{n-2} \to \mathbb{R}^{n-2} \text{ is a linear operator, } ||A|| < |s|,$$

where the z- plane is tangent to L at 0: $T_0\mathbb{C} = L$; the vector function $Q(z, y) : \mathbb{R}^n \to \mathbb{R}^{n-2}$ is analytic in $\operatorname{Re} z$, $\operatorname{Im} z$, y.

Addendum. *Let in the conditions of Proposition 3.14 the mapping ϕ depend analytically on some additional parameter. Then the corresponding coordinates (z, y) may be also chosen to depend analytically on the same parameter (then so does Q).*

Proposition 3.14 and its Addendum follow from a version of Poincaré-Dulac theorem ([2], chapter 5, section 25, subsection D), as Proposition 2.7 in Subsection 2.2.

Proof of Proposition 3.8. The commutator mappings $\phi_g : G \to G$ depend analytically on the parameter g and have common fixed point at 1. For any $g \in \Pi_{\mathbb{C},1}$ the germ of the mapping ϕ_g at 1 satisfies the conditions of Proposition 3.14 with $s = s(g)$. Let (z_g, y_g) be the corresponding local coordinates from the same proposition, chosen to depend analytically on g, see the Addendum,

$$z_g = x_{1,g} + ix_{2,g}, \ x_{1,g}, x_{2,g} \in \mathbb{R}.$$

There exists an open subset $\Pi'_{\mathbb{C},1} \subset \Pi_{\mathbb{C},1} \times G$ such that $\Pi \times 1 \subset \Pi'_{\mathbb{C},1}$ and for any $g \in \Pi_{\mathbb{C},1}$ the open set $\{h \in G \mid (g, h) \in \Pi'_{\mathbb{C},1}\}$ is mapped to itself by ϕ_g, and the coordinates (z_g, y_g) are well-defined on the latter open set. This follows from (3.20) and the Addendum. For any $g \in \Pi_{\mathbb{C},1}$ the plane $L(g) \subset \mathfrak{g}$ is tangent to the z- plane in G at 1, by definition and (3.20). For any $g \in \Pi_{\mathbb{C},1}$ consider the vector

$$\nu_g = \frac{\partial}{\partial x_{1,g}} \in L(g) : \ dz_g(\nu_g) = 1. \text{ Put } v_g(h) = z_g(h)\nu_g.$$

The vector function $v_g(h)$ satisfies the statements of Proposition 3.8. This is proved analogously to the similar statement in the proof of Proposition 2.2 in Subsection 2.2 with obvious changes. This proves Proposition 3.8. □

3.4 Nondegeneracy of the derivative. Proof of Lemma 3.11

For the proof of Lemma 3.11 we introduce the following auxiliary linear operator $\Omega = \Omega_{Y,\Sigma} : T_0\mathbb{R}^n \to \mathfrak{g}$ associated to g_j, h and l. We prove its nondegeneracy for appropriate g_j, h and l and then deduce Lemma 3.11.

Let $\Pi'_{\mathbb{C},1}$ be as in (3.5), $g_1, \dots, g_n$, h be arbitrary words in M elements, denote

$$A_j = g_j(\alpha(0)), \ H = h(\alpha(0)),$$

such that

$$(A_j, H) \in \Pi'_{\mathbb{C},1} \text{ for all } j = 1, \dots, n, \ H \text{ belongs to a 1-to-1 exponential chart.} \tag{3.21}$$

Recall that we have fixed an analytic branch of the function $s(g)$ on the set of $\mathbb{C}$-1-proximal elements $g \in G$, which contains $\Pi_{\mathbb{C},1}$, see 3.1. Denote

$$s_j(u) = s(g_j(\alpha(u))), \ \sigma_j = ds_j(0) : T_0\mathbb{R}^n \to \mathbb{C}, \ \Sigma = (\sigma_1, \dots, \sigma_n). \tag{3.22}$$

Let

$$v_{A_j} : G \to L(A_j) \subset \mathfrak{g}, \ dv_{A_j} : \mathfrak{g} \to L(A_j)$$

be the respectively the vector function from (3.6) and its differential at 1, see (3.6). The vector function v_{A_j} is well-defined in a neighborhood of unity that contains H, by (3.5) and (3.21). Let $l = (l_1, \dots, l_n) \in \mathbb{Z}^n$ be an arbitrary collection of n integers. Denote

$$\widetilde{h} = \log H \in \mathfrak{g} : \ H = \exp \widetilde{h};$$

$$Y_j(\widetilde{h}) = s^{l_j - 1}(A_j) dv_{A_j}(\widetilde{h}) \in L(A_j) \subset \mathfrak{g}, \ Y = (Y_1, \dots, Y_n), \tag{3.23}$$

$$\Omega = \Omega_{Y,\Sigma} = \sum_{j=1}^{n} Y_j\sigma_j : T_0\mathbb{R}^n \to \mathfrak{g}. \tag{3.24}$$

(The multiplication of vectors in $L(A_j)$ by complex numbers is defined in terms of the $s(A_j)$-complex structure, as in Subsection 3.1.) We show (the next proposition) that the operator $\varepsilon\Omega_{Y,\Sigma}$ is the main asymptotic term of the derivative $((\Psi(0))^{-1}\Psi)'(0)$, as $\varepsilon, \widetilde{h} \to 0$.

The principal part of the proof of Lemma 3.11 is Lemma 3.18 (stated and proved below), which says that the operator $\Omega_{Y,\Sigma}$ is nondegenerate for appropriately chosen g_j, h and l. The proof of Lemma 3.18 is done in three steps. On the first step we choose arbitrary g_j such that

$$A_j = g_j(\alpha(0)) \in \Pi_{\mathbb{C},1}, \text{ denote } \rho_j = \frac{\arg s(A_j)}{\pi}, \tag{3.25}$$

so that in addition

$$\text{the system of 1- forms } \sigma_j \text{ has real rank } n, \text{ i.e.,} \tag{3.26}$$

$$\text{the linear operator } \Sigma : T_0\mathbb{R}^n \to \mathbb{C}^n \text{ has zero kernel,}$$

$$\text{for any } j = 1,\dots,n \text{ either } \rho_j \notin \mathbb{Q}, \text{ or } \rho_j = \frac{p_j}{q_j} \in \mathbb{Q} \text{ with } q_j > n. \tag{3.27}$$

The existence of words g_j satisfying (3.25)-(3.27) easily follows from the Main Technical Lemma from Subsection 2.1.

Replacing a word g_j by its conjugate $h_j g_j h_j^{-1}$ does not change the value $s(A_j)$ and the form σ_j (the invariance of the function $s(g)$ under conjugations in G).

On the second step we show (Lemma 3.16) that replacing the words g_j by appropriate conjugates $\widetilde{g}_j = h_j g_{r_j} h_j^{-1}$ of some of them one can achieve that there exist some $Y_j \in L(A_j)$ such that the corresponding operator $\Omega_{Y,\Sigma}$ from (3.24) be nondegenerate. This is proved using only the rank condition (3.26).

On the third step, using condition (3.27) on the ratios ρ_j we show (Lemma 3.17) that the latter Y_j can be realized as $Y_j(\widetilde{h})$, see (3.23), with arbitrary given $\widetilde{h} \in \mathfrak{g}$ (independent on j) such that $Y_j(\widetilde{h}) \neq 0$ and appropriate integer collection $l \in \mathbb{Z}^n$ (depending on $\widetilde{h}$). This will prove Lemma 3.18.

The operator $\Omega_{Y,\Sigma}$ depends only on A_j, σ_j and $Y_j \in L(A_j)$ (the latter Y_j depend on $\widetilde{h}$). We formulate the previously mentioned Lemmas 3.16 and 3.17 in more generality: in terms of A_j and σ_j only, without using words g_j.

The deduction of Lemma 3.18 from the Main Technical Lemma and Lemmas 3.16, 3.17, and the deduction of Lemma 3.11 from the next proposition and Lemma 3.18 will be done at the end of the subsection.

Let

$$\nu_j = v_{A_j}(H), \Psi_j(\widetilde{u}) = \exp(\varepsilon s^{l_j}(A_j)e^{(d\ln s_j(0))\widetilde{u}}\nu_j), \tag{3.28}$$

$$\Psi(\widetilde{u}) = \Psi_1(\widetilde{u})\dots\Psi_n(\widetilde{u}), \tag{3.29}$$

see formula (3.11) in 3.1.

We identify the tangent spaces

$$T_{\Psi_j(0)}G \text{ and } T_{\Psi(0)}G \text{ with } T_1G = \mathfrak{g}$$

via left multiplication by $(\Psi_j(0))^{-1}$ and $(\Psi(0))^{-1}$ respectively. Thus, we consider the derivatives

$$\Psi_j'(0),\ \Psi'(0) \text{ as linear operators } T_0\mathbb{R}^n \to \mathfrak{g}.$$

Proposition 3.15 *Let g_j, h, $\widetilde{h}$, Σ, $l = (l_1, \dots, l_n)$, $Y = (Y_1, \dots, Y_n)$, $\Omega_{Y,\Sigma}$ be as in (3.21) - (3.24). Let $\Psi(\widetilde{u})$ be the corresponding mapping $\mathbb{R}^n \to G$ from (3.11). Let $\Psi'(0)$ be its derivative at 0 considered as a linear operator $T_0\mathbb{R}^n \to \mathfrak{g}$, see the previous paragraph. One has*

$$\Psi'(0) = \varepsilon(\Omega_{Y,\Sigma} + o(\widetilde{h})), \text{ as } \varepsilon, \widetilde{h} \to 0. \tag{3.30}$$

Proof Let $\Psi_j(\widetilde{u})$ be the mappings from (3.28). By (3.29) and (2.23), one has

$$\Psi'(0) = \Psi_n'(0) + Ad^{-1}_{\Psi_n(0)}\Psi'_{n-1}(0) + \dots + Ad^{-1}_{\Psi_2(0)\dots\Psi_n(0)}\Psi_1'(0). \text{ Thus,}$$

$$\Psi'(0) = \sum_{j=1}^n \Psi_j'(0) + \Delta,\ |\Delta| \leq n \max_{j=2,\dots,n} ||Ad^{-1}_{\Psi_j(0)\dots\Psi_n(0)} - Id|| \max_{j=1,\dots,n} |\Psi_j'(0)|. \tag{3.31}$$

One has

$$\Psi_j'(0) = \varepsilon\nu_j s^{l_j-1}(A_j)\sigma_j + o(\varepsilon\nu_j), \text{ as } \varepsilon, \widetilde{h} \to 0, \text{ by definition,}$$
$$\nu_j = v_{A_j}(\exp(\widetilde{h})) = dv_{A_j}(\widetilde{h}) + o(\widetilde{h}) = O(\widetilde{h}). \text{ Hence,}$$

$$\Psi_j'(0) = \varepsilon(dv_{A_j}(\widetilde{h}) + o(\widetilde{h}))s^{l_j-1}(A_j)\sigma_j + o(\varepsilon\widetilde{h}) = \varepsilon Y_j(\widetilde{h})\sigma_j + o(\varepsilon\widetilde{h}). \text{ One has} \tag{3.32}$$

$$\Delta = O(\varepsilon^2|\widetilde{h}|^2) \tag{3.33}$$

by (3.31), (3.32) and since the difference in the right-hand side of (3.31) is $O(\varepsilon\widetilde{h})$. The latter statement follows from the asymptotics $dist(\Psi_r(0), 1) = O(\varepsilon\widetilde{h})$, which holds true by definition. Now formulas (3.31)-(3.33) imply (3.30). Proposition 3.15 is proved. □

Lemma 3.16 *Let G, n be as at the beginning of the section. Let $A_1, \dots, A_n \in G$ be arbitrary collection of $\mathbb{C}$-1-proximal elements, $L(A_j) \subset \mathfrak{g}$ be the corresponding Ad_{A_j}- invariant planes equipped with the $s(A_j)$- complex structures (see Definition 3.6). Let $\Sigma = (\sigma_1, \dots, \sigma_n)$ be a collection of $\mathbb{R}$- linear complex-valued 1- forms $\sigma_j : \mathbb{R}^n \to \mathbb{C}$ of real rank n, i.e., $Ker\Sigma = 0$. There exists a collection of n (not necessarily distinct) indices*

$$r_1, \dots, r_n \in \{1, \dots, n\} \text{ and } n \text{ elements } H_1, \dots, H_n \in G, \text{ denote}$$
$$\widetilde{A}_j = H_j A_{r_j} H_j^{-1},\ \widetilde{\sigma}_j = \sigma_{r_j},\ \widetilde{\Sigma} = (\widetilde{\sigma}_1, \dots, \widetilde{\sigma}_n),$$

such that there exists a collection

$$Y = (Y_1, \dots, Y_n) \text{ of vectors } Y_j \in L(\widetilde{A}_j)$$

for which the linear operator

$$\Omega_{Y,\widetilde{\Sigma}} = \sum_{j=1}^n Y_j\widetilde{\sigma}_j : \mathbb{R}^n \to \mathfrak{g}$$

be nondegenerate (and hence, an isomorphism).

Proof For any $\mathbb{C}$-1-proximal element $A \in G$ and any $H \in G$ one has

$$L(HAH^{-1}) = Ad_H L(A)$$

by definition. We show that one can choose $H_1, \dots, H_n \in G$, indices $r_j = 1, \dots, n$ (some indices r_j may coincide) and vectors $Y_j \in L(\widetilde{A}_j) = Ad_{H_j} L(A_{r_j})$ so that for any $j = 1, \dots, n$ the linear operator

$$\Omega_j = \sum_{i=1}^{j} Y_i \sigma_{r_i} : \mathbb{R}^n \to \mathfrak{g} \text{ has kernel } K_j = Ker\Omega_j \text{ of codimension at least } j. \tag{3.34}$$

(By definition, $\Omega_{Y,\widetilde{\Sigma}} = \Omega_n$.) This will prove the lemma.

We construct the previous H_j, r_j, Y_j by induction in j.

Induction base: $j = 1$. Take arbitrary r_1 so that

$$\sigma_{r_1} \not\equiv 0 \text{ and arbitrary } H_1 \in G,\ Y_1 \in L(\widetilde{A}_1) \setminus 0. \text{ Then}$$

$$\Omega_1 = Y_1 \sigma_{r_1} \not\equiv 0$$

and statement (3.34) is obvious.

Induction step: $1 < j \leq n$. Let we have already chosen r_i, H_i, Y_i for $i \leq j-1$ so that (3.34) holds true with j replaced by $j-1$. Let Ω_{j-1} be the corresponding operator, $K_{j-1} = Ker\Omega_{j-1}$. Let us construct r_j, H_j and Y_j for which (3.34) holds.

By the induction hypothesis,

$$codimK_{j-1} \geq j-1.$$

If $codimK_{j-1} \geq j$, then (3.34) holds true with $Y_j = 0$ and arbitrary r_j, H_j: in this case $K_j = K_{j-1}$. Thus, everywhere below we consider that

$$dim\Omega_{j-1}(\mathbb{R}^n) = codimK_{j-1} = j-1 < n, \text{ in particular, } K_{j-1} \neq 0,\ \Omega_{j-1}(\mathbb{R}^n) \neq \mathfrak{g}. \tag{3.35}$$

Let us fix a $r_j \in \{1, \dots, n\}$ such that

$$\sigma_{r_j}|_{K_{j-1}} \not\equiv 0. \tag{3.36}$$

It exists by (3.35) and the rank condition of the lemma. Let us fix a $H_j \in G$ (denote $\widetilde{A}_j = H_j A_{r_j} H_j^{-1}$) such that

$$L = L(\widetilde{A}_j) = Ad_{H_j} L(A_{r_j}) \not\subset \Omega_{j-1}(\mathbb{R}^n). \text{ Denote } \Lambda = \Omega_{j-1}(\mathbb{R}^n) \cap L. \tag{3.37}$$

The possibility to choose H_j satisfying (3.37) follows from the irreducibility of Ad_G and (3.35). Then Λ is a linear subspace in the plane L and $\Lambda \neq L$. Thus, either $\Lambda = 0$, or Λ is a line in L. Given the previous r_j and H_j, choosing a $Y_j \in L$ defines Ω_j and hence, K_j. For any choice of $Y_j \in L$ one has

$$K_{j-1}, K_j \subset P = \Omega_{j-1}^{-1}(\Lambda) = \Omega_{j-1}^{-1}(L), \tag{3.38}$$

by definition and since

$$\Omega_j = \Omega_{j-1} + Y_j \sigma_{r_j} : \text{ thus, } \Omega_j(v) = 0 \text{ implies } \Omega_{j-1}(v) = -Y_j \sigma_{r_j}(v) \in \Omega_{j-1}(\mathbb{R}^n) \cap L = \Lambda. \tag{3.39}$$

Case 1: $\Lambda = 0$. Take arbitrary $Y_j \in L \setminus 0$. Then statement (3.34) holds true. Indeed, $\Omega_{j-1}(\mathbb{R}^n) \cap (Y_j \sigma_{r_j}(\mathbb{R}^n)) = 0$, since the latter intersection is contained in $\Omega_{j-1}(\mathbb{R}^n) \cap L = \Lambda = 0$ by definition. Hence, by (3.39),

$$K_j = Ker(\sigma_{r_j}|_{K_{j-1}}), \tag{3.40}$$

which is smaller than K_{j-1} by (3.36). Hence, $codim K_j \geq codim K_{j-1} + 1 = j$. This proves (3.34).

Case 2: Λ is a line. Consider the preimage

$$P = \Omega_{j-1}^{-1}(\Lambda) = \Omega_{j-1}^{-1}(L),$$

see (3.38). By definition,

$$codim P = codim K_{j-1} - 1 = j - 2. \text{ One has } \sigma_{r_j}(P) \neq 0 \text{ by (3.36) and (3.38).}$$

Consider the following subcases, when $\sigma_{r_j}(P) \subset \mathbb{C}$ is respectively a line in $\mathbb{C}$ or the whole $\mathbb{C}$.

Case 2.1: $\sigma_{r_j}(P) \subset \mathbb{C}$ is a line. Take a $Y_j \in L \setminus 0$ so that

$$\text{the line } \hat{l} = Y_j \sigma_{r_j}(P) \subset L \text{ be distinct from the line } \Lambda. \text{ Then}$$

$$\Omega_j = \Omega_{j-1} + Y_j \sigma_{r_j}, \ \Omega_{j-1}(P) \cap (Y_j \sigma_{r_j}(P)) = \Lambda \cap \hat{l} = 0.$$

This implies (3.40) and together with (3.36) proves (3.34), as in the previous case, when $\Lambda = 0$.

Case 2.2: $\sigma_{r_j}(P) = \mathbb{C}$. Fix an arbitrary $Y_j \in L \setminus 0$. Then

$$Y_j \sigma_{r_j}(P) = L. \text{ Put } P' = \{v \in P \mid Y_j \sigma_{r_j}(v) \in \Lambda\}. \text{ One has}$$

$$K_j \subset P' \subset P, \ P' \text{ is a hyperplane in } P, \tag{3.41}$$

by definition, (3.38) and since the equality $\Omega_j(v) = 0$ with $v \in P$ implies

$$Y_j \sigma_{r_j}(v) = -\Omega_{j-1}(v) \in \Lambda = \Omega_{j-1}(P), \text{ and hence, } v \in P'. \text{ One also has} \tag{3.42}$$

$$codim P' = codim P + 1 = j - 1, \ \sigma_{r_j}|_{P'} \not\equiv 0, \tag{3.43}$$

by (3.41) and since $\sigma_{r_j}(P) = \mathbb{C}$. Thus, the kernel K_j is the subspace in P' (see (3.41)) defined by linear equation (3.42). The right-hand side of (3.42) is independent on Y_j. Its left-hand side is $Y_j \sigma_{r_j}(v)$, where the form σ_{r_j} is also independent on Y_j and does not vanish identically on P', see (3.43). Therefore, choosing $Y_j \in L$ large enough, one can achieve that the equation (3.42) be non-identical in $v \in P'$. Then the space K_j of its solutions is smaller than P', thus, $codim K_j \geq codim P' + 1 = j$. The induction step is over. This proves (3.34). Lemma 3.16 is proved. □

Lemma 3.17 *Let G, n be as at the beginning of the section, $\Pi_{\mathbb{C},1} \subset G$ be as in (3.4). Let $A_1, \dots, A_n \in \Pi_{\mathbb{C},1}$. Let $L(A_j) \subset \mathfrak{g}$ be the corresponding Ad_{A_j}- invariant planes equipped with the $s(A_j)$- complex structures (see Definition 3.6). Let the ratios $\rho_j = \frac{\arg s(A_j)}{\pi}$ satisfy (3.27).*

Let $\Sigma = (\sigma_1, \dots, \sigma_n)$ *be a collection of* $\mathbb{R}$*- linear complex-valued 1- forms* $\sigma_j : \mathbb{R}^n \to \mathbb{C}$ *such that there exists a vector collection* $Y = (Y_1, \dots, Y_n)$*,* $Y_j \in L(A_j)$*, for which*

$$\text{the operator } \Omega_{Y,\Sigma} = \sum_{j=1}^{n} Y_j \sigma_j : \mathbb{R}^n \to \mathfrak{g} \text{ be nondegenerate.}$$

Then for any other collection

$$\widetilde{Y} = (\widetilde{Y}_1, \dots, \widetilde{Y}_n), \ \widetilde{Y}_j \in L(A_j) \setminus 0,$$

there exists an integer collection $l = (l_1, \dots, l_n) \in \mathbb{Z}^n$*, denote*

$$Y_j' = s^{l_j - 1}(A_j)\widetilde{Y}_j, \ Y' = (Y_1', \dots, Y_n'),$$

such that the operator $\Omega_{Y',\Sigma} = \sum_{j=1}^{n} Y_j' \sigma_j : \mathbb{R}^n \to \mathfrak{g}$ *be nondegenerate (i.e., an isomorphism).*

Proof Consider the set

$$Reg = \{Y = (Y_1, \dots, Y_n) \in L(A_1) \oplus \dots \oplus L(A_n) = \mathbb{R}^{2n} \mid \text{the operator } \Omega_{Y,\Sigma} \text{ is nondegenerate}\}.$$

Fix some bases in the planes $L(A_j)$. The set Reg is nonempty by the condition of the lemma. It is Zariski open and its complement is a zero set of a homogeneous polynomial of degree n in the components of Y_j in the previous bases. The latter polynomial is the determinant of the matrix of the n- dimensional operator $\Omega_{Y,\Sigma}$ (in some fixed bases in $\mathbb{R}^n$ and $\mathfrak{g}$). Both $\Omega_{Y,\Sigma}$ and the coefficients of its $(n \times n)$ matrix depend linearly on Y by definition. Thus, the previous determinant is a homogeneous polynomial of degree n.

Fix arbitrary collection $\widetilde{Y} = (\widetilde{Y}_1, \dots, \widetilde{Y}_n)$, $\widetilde{Y}_j \in L(A_j) \setminus 0$. Denote

$$L_j = \{s^q(A_j)\widetilde{Y}_j \mid q \in \mathbb{Z}\} \subset L(A_j), \ L = L_1 \times \dots \times L_n \subset \mathbb{R}^{2n}. \tag{3.44}$$

Below we show that

$$L \cap Reg \neq \emptyset. \tag{3.45}$$

Then each $Y' \in L \cap Reg$ satisfies the statements of Lemma 3.17 by definition. This will prove Lemma 3.17.

Denote

$$\overline{L}_j \subset L(A_j), \ \overline{L} \subset \mathbb{R}^{2n}$$

the Zariski closure of L_j in $L(A_j)$ and that of L in $\mathbb{R}^{2n}$ respectively,

$$I = \{j = 1, \dots, n \mid \rho_j \notin \mathbb{Q}\}, \ J = \{j = 1, \dots, n \mid \rho_j = \frac{p_j}{q_j} \in \mathbb{Q}\}; \ q_j > n \text{ by (3.27)},$$

$$\Lambda_{j,q} = \text{ the line } \mathbb{R}s^q(A_j)\widetilde{Y}_j \subset L(A_j). \text{ One has} \tag{3.46}$$

$$\overline{L}_j = L(A_j) \text{ if } j \in I; \ \overline{L}_j = \cup_{q=0}^{q_j - 1} \Lambda_{j,q} \text{ if } j \in J; \ I \cup J = \{1, \dots, n\}. \tag{3.47}$$

This follows from definition and the convergence $s^q(A_j) \to 0$, as $q \to +\infty$ ($|s(A_j)| < 1$, since $A_j \in \Pi_{\mathbb{C},1}$ by assumption). By definition,

$$\overline{L} = \overline{L}_1 \times \dots \times \overline{L}_n. \tag{3.48}$$

Thus, for any $j \in J$ the set $\overline{L}_j$ is a union of $q_j > n$ lines $\Lambda_{j,q}$. Below we show that the set $\overline{L}$ (or equivalently, L) cannot be contained in the zero set of a polynomial of degree n. In particular, $L \not\subset \mathbb{R}^{2n} \setminus Reg$, which implies (3.45) and thus, the lemma.

We prove the previous statement by contradiction. Suppose the contrary: there exists a nonzero polynomial P of degree n (in the components of the vectors $Y_j \in L(A_j)$) that vanishes identically on $\overline{L}$. We show that $P \equiv 0$ on $\mathbb{R}^{2n}$, - a contradiction to the nontriviality of P. To do this, we introduce the auxiliary subspaces $L_{J',D_{J'}} \subset \mathbb{R}^{2n}$ defined as follows. Let us take arbitrary subset and collection

$$J' \subset J, \ D_{J'} = (d_j)_{j\in J'}, \ 0 \leq d_j \leq q_j - 1. \text{ Let } \Lambda_{j,d_j} \text{ be the lines from (3.46). Put} \tag{3.49}$$

$$L_{J',D_{J'}} = (\prod_{j\in I\cup(J\setminus J')} L(A_j)) \times (\prod_{j\in J'} \Lambda_{j,d_j}), \ L_\emptyset = \mathbb{R}^{2n}. \text{ One has} \tag{3.50}$$

$$\overline{L} = \cup_{D_J} L_{J,D_J} \text{ by (3.47) and (3.48).} \tag{3.51}$$

We show that

$$P|_{L_{J',D_{J'}}} \equiv 0 \text{ for any } L_{J',D_{J'}} \text{ as in (3.50), including } L_\emptyset = \mathbb{R}^{2n}, \tag{3.52}$$

by induction in the cardinality $|J \setminus J'|$.

Induction base: $J \setminus J' = \emptyset$. Then $J' = J$ and $L_{J,D_J} \subset \overline{L}$ for any D_J as in (3.49). One has $P \equiv 0$ on $\overline{L}$ (by definition), and hence, $P \equiv 0$ on L_{J,D_J}.

Induction step: $|J \setminus J'| = k \geq 1$. We assume the induction hypothesis holds true:

$$P|_{L_{J'',D_{J''}}} \equiv 0 \text{ for any } J'', D_{J''} \text{ as in (3.49) with } |J \setminus J''| < k.$$

If $J' \neq \emptyset$, we fix arbitrary $D_{J'} = (d_j)_{j\in J'}$. Let us show that $P|_{L_{J',D_{J'}}} \equiv 0$. To do this, fix arbitrary $s \in J \setminus J'$. Let $\Lambda_{s,q}$, $q = 0, \dots, q_s - 1$, be the corresponding lines from (3.46). Put

$$J'' = J' \cup \{s\}, \ d_{j,q} = d_j \text{ for } j \in J', \ d_{s,q} = q, \ D_{J''}(q) = (d_{j,q})_{j\in J''}.$$

The linear space $L_{J',D_{J'}}$ contains q_s distinct hyperplanes $L_{J'',D_{J''}(q)}$, $q = 0, \dots, q_s - 1$. The polynomial P vanishes identically on any of the latter hyperplanes (the induction hypothesis). One has $n = degP < q_s$, see (3.27). Hence, P vanishes identically on the ambient space $L_{J',D_{J'}}$. If $J' = \emptyset$, the previous discussion applies with obvious changes and shows that $P|_{L_\emptyset} \equiv 0$. The induction step is over. Statement (3.52) is proved.

The particular case of statement (3.52) for the space $L_\emptyset = \mathbb{R}^{2n}$ says that $P|_{\mathbb{R}^{2n}} \equiv 0$, - a contradiction to the nontriviality of the polynomial P. This proves Lemma 3.17. □

Lemma 3.18 *Let G, n, M, $\alpha(u)$ be as at the beginning of the Section, $\Pi_{\mathbb{C},1} \subset G$ be the subset from (3.4). There exist words $g_1, \dots, g_n$ in M elements with $A_j = g_j(\alpha(0)) \in \Pi_{\mathbb{C},1}$ that satisfy the following statements. Let Σ be the 1- form collection (3.22), $v_{A_j} : G \to L(A_j) \subset \mathfrak{g}$ be the vector function from (3.6), $dv_{A_j} : T_0\mathbb{R}^n \to L(A_j)$ be its differential at 1, see (3.6). For any $h' \in \mathfrak{g}$ such that $dv_{A_j}(h') \neq 0$ for all j there exists an integer collection $l = (l_1, \dots, l_n) \in \mathbb{Z}^n$ such that the corresponding operator $\Omega_{Y(h'),\Sigma}$, see (3.23), (3.24), be nondegenerate.*

Proof In the proof of Lemma 3.18 we use Lemmas 3.16, 3.17, the Main Technical Lemma and the following

Claim. There exists a neighborhood $W \subset G$ of unity (covered by an exponential chart), denote

$$W' = \{\mathbb{C}-1- \text{ proximal elements in } W\}, \text{ such that}$$

$$\text{the function } \arg s(g) \text{ (which is real-analytic on } W') \text{ is locally nonconstant.} \tag{3.53}$$

Proof Given a $g_0 \in W'$, let us show that $\arg s(g) \not\equiv const$ in a neighborhood of g_0. Let $\Gamma \subset G$ be a one-parametric subgroup through g_0 (which exists since g_0 is covered by an exponential chart), $v \in \mathfrak{g}$ be its tangent vector at 1, i.e.,

$$\Gamma = \{g(\tau) = \exp(\tau v) \mid \tau \in \mathbb{R}\}.$$

The function

$$\lambda(g) = s(g) + 1 \text{ defines an eigenvalue of } Ad_g$$

The functions $s(g(\tau))$, $\lambda(g(\tau))$ extend analytically from τ_0, $g(\tau_0) = g_0$, to all the $\tau \in \mathbb{R}$ so that $\lambda(g(\tau)) : \mathbb{R} \to \mathbb{C}^*$ is a homomorphism. Thus, the value $\lambda(g(\tau))$ spirals around 0, as $\tau \to 0$, since $\operatorname{Im}\lambda(g_0) = \operatorname{Im} s(g_0) \neq 0$ (the element g_0 is $\mathbb{C}$-1-proximal). Therefore, the values $s(g(\tau)) = \lambda(g(\tau)) - 1$ cannot range in a fixed real line in $\mathbb{C}$, and hence, $\arg s(g)$ is a nonconstant analytic function in a neighborhood of g_0. The Claim is proved. □

Step 1: choice of the words g_j. Let $W' \subset G$ be the set from (3.53). The function $\arg s(g)$ is analytic on W' and nonconstant by (3.53). We apply the Main Technical Lemma to the function $\arg s(g)$ and a collection of elements

$$A'_1, \dots, A'_n \in W' \cap \Pi_{\mathbb{C},1} \text{ with } \rho'_j = \frac{\arg s(A'_j)}{\pi} \notin \mathbb{Q}.$$

The existence of the latter A'_j follows from statement (3.53) and continuity. We choose a $\delta > 0$ so that the δ- neighborhood of each A'_j lies in $\Pi_{\mathbb{C},1}$ and contains no element g with $\frac{\arg s(g)}{\pi} = \frac{p}{q} \in \mathbb{Q}$, $0 < q \leq n$. Then the corresponding words g_j given by the Main Technical Lemma whose values $A_j = g_j(\alpha(0))$ give δ- approximations of A'_j satisfy statements (3.25)-(3.27).

Step 2: making the operator $\Omega_{Y,\Sigma}$ nondegenerate for some Y. To do this, we fix some words $g_1, \dots, g_n$ constructed on Step 1. We show that one can choose an index collection $r_1, \dots, r_n \in \{1, \dots, n\}$, a word collection $h_1, \dots, h_n$, denote

$$\widetilde{g}_j = h_j g_{r_j} h_j^{-1}, \ \widetilde{A}_j = \widetilde{g}_j(\alpha(0)), \ \widetilde{\sigma}_j = ds(\widetilde{g}_j(\alpha(u))|_{u=0} : T_0\mathbb{R} \to \mathbb{C}, \ \widetilde{\Sigma} = (\widetilde{\sigma}_1, \dots, \widetilde{\sigma}_n), \tag{3.54}$$

and a $Y = (Y_1, \dots, Y_n)$, $Y_j \in L(\widetilde{A}_j)$, so that the corresponding operator $\Omega_{Y,\widetilde{\Sigma}} = \sum_{j=1}^n Y_j \widetilde{\sigma}_j$ be nondegenerate. To do this, we apply Lemma 3.16 to the elements $A_j = g_j(\alpha(0))$ and the forms

$$\sigma_j = ds(g_j(\alpha(u)))|_{u=0} : T_0\mathbb{R}^n \to \mathbb{C}.$$

Let $H_1, \dots, H_n \in G$, $r_1, \dots, r_n \in \{1, \dots, n\}$, Y be respectively the corresponding conjugating elements, indices and vector collection. Without loss of generality we consider that

$$H_j = h_j(\alpha(0)) \text{ for some words } h_j$$

(density of the subgroup $< \alpha(0) >\subset G$ and the persistence of the nondegeneracy of $\Omega_{Y,\widetilde{\Sigma}}$ under small perturbations of the elements H_j and the vectors Y_j). Fix the latter words h_j. Then by definition, the forms

$$\widetilde{\sigma}_j = \sigma_{r_j} \text{ from Lemma 3.16 coincide with the 1- forms } \widetilde{\sigma}_j \text{ from (3.54).}$$

By Lemma 3.16, the operator

$$\Omega_{Y,\widetilde{\Sigma}} = \sum_j Y_j \widetilde{\sigma}_j \text{ is nondegenerate.}$$

Step 3: making $\Omega_{Y,\widetilde{\Sigma}}$ nondegenerate with

$$Y_j = Y_j(h') = s^{l_j-1}(\widetilde{A}_j) dv_{\widetilde{A}_j}(h'), \ Y(h') = (Y_1(h'), \dots, Y_n(h')), \tag{3.55}$$

and appropriately chosen l_j and h'.

Let $\widetilde{g}_j$ be the words constructed on Step 2, $\widetilde{A}_j$, $\widetilde{\Sigma}$ be as in (3.54). Let $h' \in \mathfrak{g}$ be arbitrary vector such that

$$\widetilde{Y}_j = dv_{\widetilde{A}_j}(h') \neq 0 \text{ for any } j; \text{ denote } \widetilde{Y} = (\widetilde{Y}_1, \dots, \widetilde{Y}_n).$$

Applying Lemma 3.17 to the above $\widetilde{A}_j$, $\widetilde{\Sigma}$ and $\widetilde{Y}$ yields an integer collection $l = (l_1, \dots, l_n) \in \mathbb{Z}^n$ such that the operator $\Omega_{Y(h'),\widetilde{\Sigma}}$ with $Y(h')$ as in (3.55) is nondegenerate. Thus, the words $\widetilde{g}_j$ and the latter collection l satisfy the statements of Lemma 3.18. This proves Lemma 3.18. □

Proof of Lemma 3.11. Let $g_1, \dots . g_n$, $h' \in \mathfrak{g}$, $l = (l_1, \dots, l_n) \in \mathbb{Z}^n$ be respectively words, vector and an integer collection that satisfy the statements of Lemma 3.18: the operator $\Omega_{Y(h'),\Sigma}$, see (3.22)-(3.24), is nondegenerate. Let us show that there exists a word h such that for any $\varepsilon > 0$ small enough the derivative $\Psi'(0)$ of the corresponding mapping $\Psi(\widetilde{u})$ from (3.11) be nondegenerate. To do this, we use Proposition 3.15 and the following lower bound for $\Omega_{Y(h''),\Sigma}$: there exist a $c > 0$ and an open cone $K \subset \mathfrak{g}$, $h' \in K$, such that for any $h'' \in K$ and any $v \in T_0\mathbb{R}^n$ one has

$$||\Omega_{Y(h''),\Sigma}(v)|| \geq c||h''||||v||. \tag{3.56}$$

This follows from the linear dependence of the operator $\Omega_{Y(h''),\Sigma}(v)$ on the parameter h'' and the persistence of its nondegeneracy with $h'' = h'$ under small perturbations of h''.

Recall that by Proposition 3.15,

$$\Psi'(0) = \varepsilon(\Omega_{Y(\widetilde{h}),\Sigma} + o(\widetilde{h})), \text{ as } \varepsilon, \widetilde{h} \to 0. \tag{3.57}$$

Once the words g_j and the integer collection l are fixed, the derivative $\Psi'(0)$, and hence, the latter operator-valued function $o(\widetilde{h})$, depend only on ε and $\widetilde{h}$. Let c be as in (3.56). Fix a $\varepsilon_0 > 0$ and a neighborhood W of zero in $\mathfrak{g}$ such that for any $\varepsilon < \varepsilon_0$ and any $\widetilde{h} \in W$

$$||o(\widetilde{h})|| \leq \frac{c}{2}||\widetilde{h}|| \tag{3.58}$$

(These ε_0 and W exist by definition.) Fix an arbitrary word h such that

$$\widetilde{h} = \log(h(\alpha(0))) \in W \cap K.$$

Then for any $\varepsilon < \varepsilon_0$ and any $v \in T_0\mathbb{R}^n$ one has

$$||\Psi'(0)v|| \geq \frac{c}{2}\varepsilon||\widetilde{h}||||v||,$$

by (3.57), (3.56) and (3.58). Thus, the derivative $\Psi'(0)$ is nondegenerate for any $\varepsilon < \varepsilon_0$. The proof of Lemma 3.11 is complete. □

4 Proof of Theorem 1.1 for arbitrary Lie group

We have already proved Theorem 1.1 for any semisimple Lie group with irreducible adjoint. Let us prove it for arbitrary Lie group. To do this, we use the following

Proposition 4.1 *Let $G = H_1 \times \cdots \times H_s$ be a direct product of groups $H_1, \dots, H_s$, $\pi_j : G \to H_j$ be the corresponding projections. Let $g_1, \dots, g_s \in G$ be its elements such that $\pi_j(g_j) = 1$ for all j. Then*

$$[\dots[[g_1, g_2], g_3], \dots, g_s] = 1.$$

Proof One has $\pi_1([g_1, g_2]) = \pi_2([g_1, g_2]) = 1$, since $\pi_1(g_1) = \pi_2(g_2) = 1$. Analogously $\pi_j([[g_1, g_2], g_3]) = 1$ for any $j = 1, 2, 3$, etc. At the end we obtain that all the projections of the previous s- th commutator vanish. □

Recall that without loss of generality we assume that the elements $A, B \in G$ under consideration generate a dense subgroup.

Case 1), when G is semisimple. Let

$$\pi : G \to H_1 \times \cdots \times H_s,\ \pi_j : G \to H_j$$

be respectively the homomorphism from Proposition 1.42 and its composition with the product projection to H_j. Recall that each adjoint Ad_{H_j} is irreducible (thus, Theorem 1.1 is already proved for each H_j). The homomorphism π is a local diffeomorphism, and each π_j is surjective (Proposition 1.42). Denote

$$A_{(j)} = \pi_j(A) \in H_j,\ B_{(j)} = \pi_j(B) \in H_j.$$

The subgroup $< A_{(j)}, B_{(j)} > \subset H_j$ is dense, as is $< A, B >$ (the previous surjectivity statement). By Theorem 1.1, for any $j = 1, \dots, s$ there exist sequences $w_{(j)k}(a, b)$ of nontrivial words and pairs $(A_{(j)k}, B_{(j)k}) \in H_j \times H_j$, $(A_{(j)k}, B_{(j)k}) \to (A_{(j)}, B_{(j)})$, such that $w_{(j)k}(A_{(j)k}, B_{(j)k}) = 1$. The latter pair sequences can be lifted up to a sequence $(A_k, B_k) \to (A, B)$ such that

$$\pi_j(A_k) = A_{(j)k},\ \pi_j(B_k) = B_{(j)k}$$

(by the local diffeomorphicity of π). We show that there are nontrivial relations $w_k(A_k, B_k) = 1$ in the groups $< A_k, B_k >$. This will prove Theorem 1.1. Firstly put

$$w_k = [\dots[[w_{(1)k}, w_{(2)k}], w_{(3)k}], \dots, w_{(s)k}]. \text{ One has } \pi(w_k(A_k, B_k)) = 1, \tag{4.1}$$

by definition and Proposition 4.1 applied to $g_j = \pi(w_{(j)k}(A_k, B_k))$.

Case A): the commutators w_k are nontrivial words (for any k large enough).

Subcase (Ai): the homomorphism π is injective. Then $w_k(A_k, B_k) = 1$ by (4.1). Thus, the words w_k and the pairs (A_k, B_k) satisfy the statements of Theorem 1.1. This proves Theorem 1.1.

Subcase (Aii): the homomorphism π is not injective. Consider its kernel K, which is a subgroup in the center $C(G_0)$ of the unity component G_0 of G (Proposition 1.42). Relations (4.1) imply that $w_k(A_k, B_k) \in K$. The subgroup $K \subset G$ is normal. Therefore, the values at (A_k, B_k) of the commutators $[a, w_k(a, b)]$, $[b, w_k(a, b)]$ belong to the commutative group K, as does $w_k(A_k, B_k)$. For any k fix a symbol $d_k = a, b$ such that $w_k(a, b) \neq d_k^p$ for any $p \in \mathbb{Z}$. Then the abstract words w_k and d_k generate a free group, in particular,

$$\text{the words } \widetilde{w}_k(a, b) = [w_k(a, b), [d_k, w_k(a, b)]] \text{ are nontrivial and} \tag{4.2}$$

$$\widetilde{w}_k(A_k, B_k) = 1;\ |\widetilde{w}_k| \leq 16|w_k| \leq 4^{s+1} \max_j |w_{(j)k}|. \tag{4.3}$$

This proves Theorem 1.1. (The latter inequality will be used in the proof of Theorem 1.29.)

Case B): some commutators w_k are trivial words. We show that one can construct the previous words $w_{(j)k}$ and pairs $(A_{(j)k}, B_{(j)k})$ so that the corresponding commutators w_k be nontrivial for all k large enough. This will reduce the proof of Theorem 1.1 to the previous case A). To do this, we use Theorem 3.12 and the following

Proposition 4.2 *Let $M, s \in \mathbb{N}$, $\gamma_1, \dots, \gamma_s$ be arbitrary nontrivial words in M elements. Let η_1, η_2 be arbitrary pair of noncommuting words, put $\eta_0 = 1$. There exists a collection of indices*

$$i(s) \in \{0, 1, 2\},\ \text{put } \gamma'_j = \eta_{i(j)}^{-1}\gamma_j \text{ for any } j = 1, \dots, s,\ \text{such that}$$

$$\text{the word } []_j = [\dots[[\gamma'_1, \gamma'_2], \gamma'_3], \dots, \gamma'_j] \text{ be nontrivial for any } j = 1, \dots, s.$$

Proof We choose the indices $i(j)$ by induction in j.

Induction base. We put $i(1) = 0$: then $[]_1 = \gamma_1$ is a nontrivial word.

Induction step. Let we have already chosen indices $i(q)$ for all $q \leq j$ so that $[]_j$ be a nontrivial word. Let us show that one can choose $i(j+1) \in \{0, 1, 2\}$ so that $[]_{j+1} = [[]_j, \gamma'_{j+1}]$ be nontrivial. Indeed, suppose the contrary: then the three words γ_{j+1}, $\eta_1^{-1}\gamma_{j+1}$, $\eta_2^{-1}\gamma_{j+1}$ commute with a nontrivial word $[]_j$. Therefore, the group generated by the four previous words (which contains η_1 and η_2) is contained in a cyclic word group (liberty of a subgroup of free group). Hence, η_1 and η_2 commute, - a contradiction. □

For any $j = 1, \dots, s$ let us choose a *conj*- nondegenerate family

$$\alpha_j(u_j) = (a(u_j), b(u_j)) \in H_j \times H_j,\ \alpha_j(0) = (A_{(j)}, B_{(j)}),$$

$$u_j \in \mathbb{R}^{d_j},\ d_j = dim H_j,$$

and fix a corresponding converging tuple $(\{w'_{(j)r}\}, \{k_{jr}\}, \psi_j, \delta_j)$, see Definition 3.13:

$$k_{jr} \to \infty,\ \text{as } r \to \infty,\ \psi_j : \mathbb{R}^{d_j} \to H_j \text{ is a smooth mapping,}$$

$$\psi_j : \overline{D_{\delta_j}} \mapsto \psi(\overline{D_{\delta_j}}) \subset H_j \text{ is a diffeomorphism, } 1 \in \psi_j(D_{\delta_j}),\ w'_{(j)r}(\alpha_j(k_{jr}^{-1}\widetilde{u}_j)) \to \psi_j(\widetilde{u}_j) \tag{4.4}$$

uniformly with derivatives on compact sets in $\mathbb{R}^{d_j}$. Fix also two auxiliary words

$$\eta_1, \eta_2,\ [\eta_1, \eta_2](A, B) \neq 1,\ \pi_j(\eta_i(A, B)) \in \psi_j(D_{\delta_j}) \text{ for all } i = 1, 2,\ j = 1, \dots, s;\ \text{put } \eta_0 = 1. \tag{4.5}$$

The existence of these η_1, η_2 follows from the density of the subgroup $< A, B > \subset G$, the local diffeomorphicity of the mapping π and the noncommutativity of the unity component G_0 of G. There exists a sequence

$$i(j, r) \in \{0, 1, 2\} \text{ depending on } j = 1, \dots, s,\ r \in \mathbb{N},\ \text{put } w_{(j)r} = \eta^{-1}_{i(j,r)} w'_{(j)r},\ \text{such that} \tag{4.6}$$

$$\text{the corresponding commutators } w_r \text{ from (4.1) be nontrivial words}$$

(Proposition 4.2 and the noncommutativity of η_1 and η_2, see (4.5)). Fix the previous sequence $i(j, r)$. There exist s sequences $u_{jr} \in \mathbb{R}^{d_j}$, $u_{jr} \to 0$, as $r \to \infty$, such that $w_{(j)r}(\alpha_j(u_{jr})) = 1$. Indeed, each subsequence of the mappings

$$\chi_{j,r} : \widetilde{u}_j \mapsto w_{(j)r}(\alpha_j(k^{-1}_{jr} \widetilde{u}_j))$$

contains a smaller subsequence converging to one of the three possible limits

$$\psi_{i,j}(\widetilde{u}_j) = \eta^{-1}_i(A_{(j)}, B_{(j)}) \psi_j(\widetilde{u}_j),\ i = 0, 1, 2,$$

uniformly with derivatives on compact subsets in $\mathbb{R}^{d_j}$, by (4.4). One has $1 \in \psi_{i,j}(D_{\delta_j})$, by definition and the inclusions in (4.4) and (4.5). Therefore, for any r large enough there exists a $\widetilde{u}_{jr} \in D_{\delta_j}$, put

$$u_{jr} = k^{-1}_{jr} \widetilde{u}_{jr},\ (A_{(j)r}, B_{(j)r}) = \alpha(u_{jr}),\ \text{such that } \chi_{j,r}(\widetilde{u}_{jr}) = w_{(j)r}(A_{(j)r}, B_{(j)r}) = 1. \tag{4.7}$$

Fir a sequence of pairs $(A_r, B_r) \to (A, B)$ such that

$$(A_{(j)r}, B_{(j)r}) = (\pi_j(A_r), \pi_j(B_r)) \text{ for any } j = 1, \dots, s \text{ and } r \in \mathbb{N};\ \text{then } \pi_j(w_{(j)r}(A_r, B_r)) = 1.$$

Their existence follows from the local diffeomorphicity of π. Let w_r be the commutators (4.1) defined by the new words $w_{(j)r}$. One has

$$\pi(w_r(A_r, B_r)) = 1,$$

by the previous equalities and Proposition 4.1. The words w_r are nontrivial by construction. This reduces us to the case A) and proves Theorem 1.1 in the case, when G is semisimple.

Case 2): G is not semisimple. This means that G contains a maximal connected solvable normal Lie subgroup (denote it H) of positive dimension. The quotient $G_{ss} = G/H$ is a semisimple Lie group.

Proposition 4.3 *Let G, H, G_{ss} be as above. Let s be the length of the commutant chain in H:*

$$H^{(0)} = H,\ H^{(1)} = [H, H],\ H^{(2)} = [H^{(1)}, H^{(1)}], \dots, H^{(s)} = 1;\ H^{(s-1)} \neq 1. \tag{4.8}$$

For any nontrivial word $w(a, b)$ in two symbols and for any elements

$$(\widetilde{A}, \widetilde{B}) \in G \times G \text{ such that } w(\widetilde{A}, \widetilde{B}) \in H$$

there exists a nontrivial word $\widetilde{w}(a, b)$,

$$|\widetilde{w}| \leq 4^{s+1} |w|,\ \text{such that } \widetilde{w}(\widetilde{A}, \widetilde{B}) = 1. \tag{4.9}$$

Proof The word $w(a,b)$ does not commute with some of symbols a, b, say, with a. Put

$$\sigma_{0,0} = w, \ \sigma_{0,1} = [a, w],$$

which are nontrivial and noncommuting words, and define recurrently

$$\sigma_{i+1,0} = [\sigma_{i,0}, \sigma_{i,1}], \ \sigma_{i+1,1} = [\sigma_{i,0}, \sigma_{i,1}^{-1}]; \text{ put } \widetilde{w} = \sigma_{s,1}. \tag{4.10}$$

The word $\widetilde{w}$ is nontrivial and $\widetilde{w}(\widetilde{A}, \widetilde{B}) = 1$ by construction ($\sigma_{i,j}(\widetilde{A}, \widetilde{B}) \in H^{(i)}$ for any i, j by definition). Inequality (4.9) follows from definition. Proposition 4.3 is proved. □

Let $p : G \to G_{ss}$ be the quotient projection, $(A, B) \in G \times G$ be an irrational pair, $(A_k, B_k) \to (A, B)$, w_k be respectively sequences of pairs and nontrivial words such that $w_k(A_k, B_k) \in H$, i.e., $p(w_k(A_k, B_k)) = 1$. (They exist by Theorem 1.1 applied to G_{ss}, which was proved above.) Let $\widetilde{w}_k$ be the corresponding words from Proposition 4.3. They are nontrivial and $\widetilde{w}_k(A_k, B_k) = 1$ by definition. The proof of Theorem 1.1 is complete.

5 A short proof of Theorem 1.1 for dense subgroups in $G = PSL_2(\mathbb{C})$

Let $A, B \in PSL_2(\mathbb{C})$ generate a free dense subgroup. We prove Theorem 1.1 by contradiction. Suppose there is a (simply connected) neighborhood $V \subset PSL_2(\mathbb{C}) \times PSL_2(\mathbb{C})$ of the pair (A, B) such that each pair $(a, b) \in V$ generates a free subgroup. Thus, each word $w(a, b)$ is a holomorphic function in $(a, b) \in PSL_2(\mathbb{C}) \times PSL_2(\mathbb{C})$ with values in $PSL_2(\mathbb{C})$; distinct words define holomorphic functions with disjoint graphs over V. Using holomorphic motion of the fixed points of the elements $w(a, b) \in PSL_2(\mathbb{C})$, we construct a nonstandard measurable almost complex structure on $\overline{\mathbb{C}}$ invariant under the action of $< A, B >$ (and hence, under the action of the whole group $PSL_2(\mathbb{C})$ by density). But the only measurable almost complex structure preserved under the action of $PSL_2(\mathbb{C})$ on $\overline{\mathbb{C}}$ is the standard complex structure - a contradiction.

Remark 5.1 The author's initial proof of Theorem 1.1 in the case, when $G = PSL_2(\mathbb{C})$, followed a similar scheme (using the holomorphic motion of fixed points) but was longer than the one presented below. The final quasiconformal mapping and invariance argument, which simplified the proof essentially, is due to Étienne Ghys.

Recall that an element $b \in PSL_2(\mathbb{C})$ is called *elliptic*, if its action on $\overline{\mathbb{C}}$ is conjugated to a rotation. It is called *hyperbolic*, if it has two fixed points: one attracting and the other one repelling. Otherwise it is *parabolic*, i.e., has a unique fixed point and is conjugated to the translation. If b has two fixed points, then their multipliers are inverse to each other. The half-sum of their multipliers (denoted $\nu(b)$) is a holomorphic function $\nu : PSL_2(\mathbb{C}) \to \mathbb{C}$.

Proposition 5.2 *Let $V \subset PSL_2(\mathbb{C}) \times PSL_2(\mathbb{C})$ be an open set such that each pair $(a, b) \in V$ generates a free subgroup in $PSL_2(\mathbb{C})$. Then each element of the latter subgroup is hyperbolic.*

Proof Suppose the contrary: there exists a pair $(a, b) \in V$ and a nontrivial word w such that the multiplier of the transformation $w(a, b)$ at some its fixed point has unit modulus. This is equivalent to say that $\nu(w(a, b)) \in [-1, 1]$. There exists a pair $(c, d) \in PSL_2(\mathbb{C}) \times$

$PSL_2(\mathbb{C})$ arbitrarily close to (a,b) (in particular, lying in V) such that the multiplier of $w(c,d)$ at some its fixed point be a root of unity, or equivalently, $\nu(w(c,d)) = \cos\theta$, $\theta \in \pi\mathbb{Q}$. This follows from the nonconstance of the holomorphic function $(c,d) \mapsto \nu(w(c,d))$ and openness of holomorphic mappings. (The function $\nu(w(c,d))$ is nonconstant on $PSL_2(\mathbb{C}) \times PSL_2(\mathbb{C})$, since $w(1,1) = 1$ and the value of the word w on the generators of a Schottky group is hyperbolic.) By construction, the transformation $w(c,d)$ is elliptic of finite order, - a contradiction to the liberty of the group $< c,d >$. The proposition is proved. □

Thus, each element $w(a,b) \in PSL_2(\mathbb{C})$, $(a,b) \in V$, is hyperbolic, hence, its fixed points are analytic functions in $(a,b) \in V$. The graphs of the fixed point functions are disjoint. Indeed, otherwise, if two distinct hyperbolic elements of $PSL_2(\mathbb{C})$ have one common fixed point, then their commutator is parabolic: the latter fixed point is its unique fixed point. This contradicts the hyperbolicity of the commutator. If two hyperbolic elements have two common fixed points, then they commute, - a contradiction to the liberty.

For any $(a,b) \in V$ denote $Fix(a,b) \subset \overline{\mathbb{C}}$ the set of fixed points of all the elements of the group $< a,b >$. The set $Fix(A,B)$ is dense in $\overline{\mathbb{C}}$, since the subgroup $< A,B >$ is dense. The previous disjoint graphs of fixed point functons form a holomorphic motion over V of the sets $Fix(a,b)$, $(a,b) \in V$. They can be extended up to a global holomorphic motion: filling the whole product $V \times \overline{\mathbb{C}}$ by a union of disjoint graphs of holomorphic functions $V \to \overline{\mathbb{C}}$. This follows immediately from the density of $Fix(A,B)$ by the disjointness and elementary normal family argument (e.g., a version of Montel's theorem, see [16]).

Remark 5.3 The well-known Slodkowski theorem [18] says that any holomorphic motion in $D \times \overline{\mathbb{C}}$ of any subset of the Riemann sphere over unit disk D extends up to a holomorphic motion of the whole Riemann sphere. Here we do not use this theorem in full generality.

For any $(a,b) \in V$ denote $h_{a,b} : \overline{\mathbb{C}} \to \overline{\mathbb{C}}$ the mapping of the $\overline{\mathbb{C}}$- fiber $(a,b) \times \overline{\mathbb{C}} \subset V \times \overline{\mathbb{C}}$ to the fiber $(A,B) \times \overline{\mathbb{C}}$ defined by the holonomy of the previous holomorphic motion. In more detail, take any path in V from (a,b) to (A,B) and lift it to each one of the previous disjoint graphs in $V \times \overline{\mathbb{C}}$. By definition, the mapping $h_{a,b}$ sends the starting point of a lifted path to its end-point. The mapping $h_{a,b}$ does not depend on the choice of path by simple connectivity of V. It is a quasiconformal homeomorphism: any holomorphic motion has a quasiconformal holonomy [19]. The homeomorphism $h_{a,b}$ conjugates the actions on $\overline{\mathbb{C}}$ of the groups $< A,B >$ and $< a,b >$, since it conjugates them on the dense invariant subsets $Fix(A,B)$ and $Fix(a,b)$ in $\overline{\mathbb{C}}$, by construction. The quasiconformal homeomorphism $h_{a,b}$ transforms the standard complex structure on $\overline{\mathbb{C}}$ to a measurable almost complex structure (denoted by $\sigma(a,b)$). The latter structure is invariant under the action of the group $< A,B >$ (by definition and the previous conjugacy statement), and hence, under $PSL_2(\mathbb{C})$, by density. Now to prove the theorem, it suffices to show that for a generic pair (a,b) the almost complex structure $\sigma(a,b)$ is not standard.

For any $(a,b) \in V$ the elements a and b are hyperbolic with distinct fixed points; the latters form a quadruple denoted $Q(a,b)$ of points in $\overline{\mathbb{C}}$. If the cross-ratios of two quadruples $Q(a,b)$ and $Q(A,B)$ are distinct, then the quasiconformal homeomorphism $h_{a,b}$, which sends $Q(a,b)$ to $Q(A,B)$, is not conformal; hence, $\sigma(a,b)$ is not standard. This together with the discussion at the beginning of the section proves Theorem 1.1.

6 Approximations by pairs with relations. Proof of Theorem 1.29

Theorem 1.29 will be proved in Subsections 6.1 (case of semisimple (not necessarily connected) Lie group with irreducible adjoint), 6.2 (case of arbitrary semisimple Lie group) and 6.3 (case of non-semisimple Lie group).

6.1 Case of semisimple Lie group G with irreducible adjoint

The group $G = G_{ss}$ is semisimple. Let us fix a left-invariant metric on G.

Let $(A, B) \in G\times G$ be an irrational pair. We have to show that it is $\varepsilon(x)$- approximable by pairs with relations. The family of all the pairs (a, b) in $G\times G$ is *conj*- nondegenerate at (A, B) (Corollary 1.47). Hence, we can choose a *conj*- nondegenerate subfamily $\alpha(u) = (a(u), b(u))$ depending on a parameter $u \in \mathbb{R}^n$, $\alpha(0) = (A, B)$. Let us fix this subfamily. We construct sequences of

$$\text{nontrivial words } w'_m, \text{ parameter values } u_m, \text{ put } (A_m, B_m) = \alpha(u_m), \text{ and numbers } l'_m \in \mathbb{N}$$

(all being defined for m no less than some $q \in \mathbb{N}$) such that

$$w'_m(A_m, B_m) = 1, \ |w'_m| \leq l'_m, \tag{6.1}$$

$$dist((A_m, B_m), (A, B)) < \varepsilon(c' l'_m), \ c' = c'(A, B) \text{ is independent on } m. \tag{6.2}$$

This means that the pair (A, B) is $\varepsilon(x)$- approximable by pairs with relations (Definition 1.27). This will prove Theorem 1.29. At the end of the Subsection we prove the Addendum to Theorem 1.29 by comparing l'_m with the length majorants l_m for the $\varepsilon(x)$- approximations by words in (A, B) on $D_1 \subset G$.

To construct the words w'_m, we first fix a converging tuple $(\{w_r\}, \{k_r\}, \psi, \delta)$ associated to $\alpha(u)$ (see Definition 3.13):

$$\psi : \overline{D}_\delta \to \psi(\overline{D}_\delta) \subset G \text{ is a diffeomorphism}, \ k_r \to \infty, \ w_r(\alpha(k_r^{-1}\widetilde{u})) \to \psi(\widetilde{u}) \tag{6.3}$$

uniformly with derivatives on $\overline{D}_\delta$. There exists a $R > 0$ (let us fix it) such that

$$w_r^{-1}(A, B) \in D_R \subset G \text{ for any } r \in \mathbb{N}. \tag{6.4}$$

This follows from (6.3) (recall that $w_r(A, B) = w_r(\alpha(0)) \to \psi(0)$). By definition, the group G is $\varepsilon(x)$- approximable on $D_R \subset G$ by words in (A, B) with bounded derivatives. Let

$$\Omega_m = \Omega_{m,D_R}, \ \widetilde{l}_m = l_m(A, B, D_R), c = c(A, B, D_R) \tag{6.5}$$

be respectively the corresponding word collection and majorant sequences and the constant from Definition 1.11. For any $r \in \mathbb{N}$ fix a word

$$\omega_{rm} \in \Omega_m \text{ such that } dist(\omega_{rm}(A, B), w_r^{-1}(A, B)) < \varepsilon(c\widetilde{l}_m), \tag{6.6}$$

which exists since the set $\Omega_m(A, B)$ contains a $\varepsilon(c\widetilde{l}_m)$- net on D_R by definition. Put

$$w'_m = w_r \omega_{rm}. \tag{6.7}$$

We show that for any fixed r large enough the words w'_m are nontrivial and satisfy (6.1) and (6.2). To do this, we consider the auxiliary mappings

$$\psi_{rm}(\widetilde{u}) = w'_m(\alpha(k_r^{-1}\widetilde{u})). \text{ One has } \psi_{rm}(\widetilde{u}) \to \widetilde{\psi}(\widetilde{u}) = \psi(\widetilde{u})(\psi(0))^{-1}, \text{ as } r, m \to \infty, \quad (6.8)$$

uniformly with derivatives on $\overline{D}_\delta$. Indeed,

$$\psi_{rm}(\widetilde{u}) = w_r(\alpha(k_r^{-1}\widetilde{u}))\omega_{rm}(\alpha(k_r^{-1}\widetilde{u})), \ w_r(\alpha(k_r^{-1}\widetilde{u})) \to \psi(\widetilde{u}) \text{ by (6.3)}. \quad (6.9)$$

Let us show that

$$\omega_{rm}(\alpha(k_r^{-1}\widetilde{u})) \to (\psi(0))^{-1}. \quad (6.10)$$

(All the previous convergences are uniform with derivatives on $\overline{D}_\delta$.)

$$\text{The derivatives of the mappings } \widetilde{u} \mapsto \omega_{rm}(\alpha(k_r^{-1}\widetilde{u})) \text{ tend to zero uniformly on } \overline{D}_\delta, \quad (6.11)$$

since $k_r \to \infty$ and the derivatives of the mappings $u \mapsto \omega_{rm}(\alpha(u))$ are uniformly bounded on a neighborhood of 0 independent on r and m (by the $\varepsilon(x)$- approximability with bounded derivatives). One has

$$dist(\omega_{rm}(\alpha(k_r^{-1}\widetilde{u})), w_r^{-1}(A,B)) \to 0 \quad (6.12)$$

uniformly on $\overline{D}_\delta$, by (6.11), (6.6) and since $\varepsilon(c\widetilde{l}_m) \to 0$. Now statement (6.10) follows from (6.12), (6.11) and the convergence $w_r(A,B) \to \psi(0)$ (which holds true by (6.3)). Statements (6.9) and (6.10) together imply (6.8).

Fix a $r \in \mathbb{N}$ (large enough) such that there exists a constant $K > 0$ such that for any m large enough (dependently on r)

$$\psi_{rm} : \overline{D}_\delta \to \psi_{rm}(\overline{D}_\delta) \subset G \text{ is a diffeomorphism and } 1 = \widetilde{\psi}(0) \in \psi_{rm}(D_\delta), \quad (6.13)$$

$$||(\psi'_{rm}(x))^{-1}|| < K \text{ for any } x \in \overline{D}_\delta. \quad (6.14)$$

The existence of the previous r follows from (6.8). Put

$$u_m = k_r^{-1}\psi_{rm}^{-1}(1), \ (A_m, B_m) = \alpha(u_m). \text{ By definition, } w'_m(A_m, B_m) = 1,$$

$$|w'_m| \leq l''_m = |w_r| + \widetilde{l}_m, \quad (6.15)$$

by the inequality $|\omega_{rm}| \leq \widetilde{l}_m$, which holds true since $\omega_{rm} \in \Omega_{m,D_R}$, see Definition 1.11,

$$dist((A_m, B_m),(A,B)) = O(u_m) = O(dist(\psi_{rm}(0),1)) = O(dist(w'_m(A,B),1)) = O(\varepsilon(c\widetilde{l}_m)),$$

by (6.14) and (6.6). Thus, there exists a constant $C > 1$ such that

$$dist((A_m, B_m),(A,B)) < C\varepsilon(c\widetilde{l}_m) \quad (6.16)$$

for any m large enough (that is, for which the previous pair (A_m, B_m) is well-defined).

Let us prove (6.1) and (6.2) with the majorants l'_m defined as follows. Let $l_m = l_m(A, B, D_1)$ be the majorants for the $\varepsilon(x)$- approximations of G on $D_1 \subset G$. By (1.5), the initial majorants $\widetilde{l}_m = l_m(A, B, D_R)$ can be taken as follows:

$$\widetilde{l}_m = l + l_m, \ l = l(R) \geq 0 \text{ is independent on } m. \text{ Then} \quad (6.17)$$

$$l''_m = \Delta + l_m, \ \Delta = l + |w_r| \text{ is independent on } m. \text{ Put}$$

$$c'' = \sup_m \frac{l''_m}{l_m}, \ l'_m = c'' l_m \geq l''_m. \text{ By definition, } |w'_m| \leq l'_m, \tag{6.18}$$

see (6.15). This proves (6.1). One has

$$C\varepsilon(c\widetilde{l_m}) \leq C\varepsilon(cl_m) = C\varepsilon(c(c'')^{-1}l'_m) < \varepsilon(c'l'_m), \ c' = c(Cc'')^{-1},$$

by (6.17) and (1.1). This together with (6.16) proves (6.2). The nontriviality of the words w'_m follows from the diffeomorphicity of the corresponding mapping ψ_{rm}, see (6.13). This proves Theorem 1.29. The statement of its Addendum follows from (6.18).

6.2 Case of arbitrary semisimple Lie group

Let G be arbitrary semisimple Lie group,

$$\pi : G \to H_1 \times \cdots \times H_s$$

be the homomorphism from Proposition 1.42. Theorem 1.29 is already proved for the groups H_j (their adjoints are irreducible).

Fix some left-invariant Riemann metrics on the groups H_j. The π- pullback of their product yields a left-invariant metric on G.

Let $(A, B) \in G \times G$ be an irrational pair. For the proof of Theorem 1.29 we have to construct sequences of nontrivial words w'_m, pairs $(A_m, B_m) \to (A, B)$ and majorants l_m satisfying (6.1) and (6.2). Let

$$\pi(A) = (A_{(1)}, \dots, A_{(s)}), \ \pi(B) = (B_{(1)}, \dots, B_{(s)}), \ A_{(j)} = \pi_j(A), \ B_{(j)} = \pi_j(B) \in H_j.$$

(Recall that $\pi_j : G \to H_j$ is the compotision of π with the product projection to H_j, see Section 4.) The projection π_j is surjective (Proposition 1.42). By the conditions of Theorem 1.29, the group G is $\varepsilon(x)$- approximable by words in (A, B) with bounded derivatives. Let

$$\Omega_m = \Omega_{m,D_1}, \ l_m = l_m(A, B, D_1), c = c(A, B, D_1) \tag{6.19}$$

be respectively the word collection and majorant sequences and the constant corresponding to the $\varepsilon(x)$- approximations of G on the unit ball $D_1 \subset G$ (see Definition 1.11). Then the group H_j is $\varepsilon(x)$- approximable by words in $(A_{(j)}, B_{(j)})$ on the unit ball $D_1 \subset H_j$ with bounded derivatives, with respect to the previous word collection and majorant sequences Ω_m, l_m and the constant c. This follows from the previous similar property of the group G and the next properties of the projection π_j: it does not increase distances and maps the unit ball $D_1 \subset G$ onto the unit ball $D_1 \subset H_j$ (by the definition of the metric on G; in particular, a δ- net on the ball $D_1 \subset G$ is projected onto a δ- net on the ball $D_1 \subset H_j$).

For any $j = 1, \dots, s$ there exist sequences of nontrivial words $w_{(j)m}$ and pairs $(A_{(j)m}, B_{(j)m}) \to (A_{(j)}, B_{(j)})$ in $H_j \times H_j$ and constants $c_j, c''_j > 0$ such that for any $m \in \mathbb{N}$

$$w_{(j)m}(A_{(j)m}, B_{(j)m}) = 1, \ |w_{(j)m}| \leq l_{jm} = c''_j l_m, \tag{6.20}$$

$$dist((A_{(j)m}, B_{(j)m}), (A_{(j)}, B_{(j)})) < \varepsilon(c_j l_{jm}) = \varepsilon(c_j c''_j l_m) \tag{6.21}$$

(Theorem 1.29 and its Addendum applied to the groups H_j). There exists a sequence $(A_m, B_m) \to (A, B)$ in $G \times G$ such that

$$\pi(A_m) = (A_{(1)m}, \dots, A_{(s)m}),\ \pi(B_m) = (B_{(1)m}, \dots, B_{(s)m}),$$

since π is a local diffeomorphism. Denote

$$w_m = [\dots[[w_{(1)m}, w_{(2)m}], w_{(3)m}], \dots, w_{(s)m}]. \tag{6.22}$$

Case 1): the words w_m are nontrivial. Let $\widetilde{w}_m$ be the corresponding (nontrivial) words from (4.2). One has

$$\widetilde{w}_m(A_m, B_m) = 1,\ |\widetilde{w}_m| \leq l'_m = \hat{c} l_m,\ \hat{c} = 4^{s+1} \max_j c''_j, \tag{6.23}$$

by definition, (4.1) (Proposition 4.1) and (4.3). The words $\widetilde{w}_m$ and the pairs (A_m, B_m) satisfy (6.1) and (6.2). Indeed, statement (6.1) follows from (6.23). One has

$$dist((A_m, B_m), (A, B)) \leq \sum_{j=1}^{s} dist((A_{(j)m}, B_{(j)m}), (A_{(j)}, B_{(j)}))$$

$$< \sum_{j=1}^{s} \varepsilon(c_j c''_j l_m) < \varepsilon(c' l'_m),\ c' = (\hat{c}s)^{-1} \min_j (c_j c''_j).$$

The latter inequality follows from the definition of the metric on G, (6.21) and (1.1). This proves (6.2) and Theorem 1.29. The statement of its Addendum follows from (6.23).

Case 2): the words w_m are trivial for arbitrarily large m. For any $j = 1, \dots, s$ we choose some appropriate sequence of words $w'_m = w'_{(j)m}$, as in the previous Subsection (with G replaced by H_j, (A, B) replaced by $(A_{(j)}, B_{(j)})$), which satisfy statements (6.20) and (6.21) (as it was proved at the same place). We construct two auxiliary word sequences

$$\eta_{1,m}, \eta_{2,m} \in \Omega_m,\ [\eta_{1,m}, \eta_{2,m}] \neq 1, \text{ and replace some } w'_{(j)m} \text{ by } w_{(j)m} = \eta^{-1}_{i(j,m),m} w'_{(j)m}$$

with appropriately chosen $i = i(j, m)$ so that the commutators w_m (see (6.22)) of the new words $w_{(j)m}$ be nontrivial (this is possible by Proposition 4.2). Using results of the previous Subsection, we show that there exists a pair sequence $(A_{(j)m}, B_{(j)m}) \to (A_{(j)}, B_{(j)})$ such that statements (6.20), (6.21) hold true. This reduces us to the previous case 1) and proves Theorem 1.29 and its Addendum.

To construct the noncommuting word sequences $\eta_{1,m}$ and $\eta_{2,m}$, we use the following

Proposition 6.1 *Let G be a connected noncommutative Lie group. There exists a $C > 1$ such that for any $\delta > 0$ small enough any δ- net on $D_{C\delta} \subset G$ contains a pair of noncommuting elements.*

Proof We prove the proposition by contradiction. Suppose the contrary: there exist sequences $\delta_k \to 0$ and $C_k \to +\infty$ and a sequence $\hat{\Omega}_k$ of δ_k- nets on $D_{C_k \delta_k}$ such that for any k each two elements of $\hat{\Omega}_k$ commute. Let us show that the Lie algebra $\mathfrak{g}$ is commutative. This contradicts the conditions of the proposition and will prove the latter. By definition, $\hat{\Omega}_k \subset D_{(C_k+1)\delta_k}$. Without loss of generality we assume that $C_k \delta_k \to 0$ (one can achieve this

by shrinking C_k); then the ball $D_{(C_k+1)\delta_k}$, and hence, $\hat{\Omega}_k$ are contracted to 1, as $k \to \infty$. In particular, the set $\hat{\Omega}_k$ is contained in the 1-to-1 image of the exponential chart, whenever k is large enough. The lines in $\mathfrak{g}$ generated by the logarithms of the elements of $\hat{\Omega}_k$ become arbitrarily dense in $\mathfrak{g}$, as $k \to \infty$ (by definition and since $C_k \to \infty$). This means that given a line $l \in \mathfrak{g}$, one can find a sequence $a_k \in \hat{\Omega}_k$, $a_k \to 1$, such that the lines generated by the logarithms $\log a_k \in \mathfrak{g}$ converge to l. Let us fix two lines $l_1, l_2 \subset \mathfrak{g}$ and the latter corresponding sequences $a_k, b_k \in \hat{\Omega}_k$. The lines l_1 and l_2 commute. This follows from definition and the fact that $[\log a_k, \log b_k] = o(|\log a_k||\log b_k|)$, as $k \to \infty$ (since a_k and b_k commute and tend to 1). This proves the commutativity of $\mathfrak{g}$ and finishes the proof of the Proposition. $\square$

The construction of the words $w'_{(j)m}$. For any $j = 1, \dots, s$ fix an arbitrary *conj*-nondegenerate family

$$\alpha_j(u_j) = (a_j(u_j), b_j(u_j)) \in H_j \times H_j, \ u_j \in \mathbb{R}^{d_j}, \ d_j = dimH_j, \ \alpha_j(0) = (A_{(j)}, B_{(j)}),$$

and a corresponding converging tuple (see Definition 3.13)

$$(\{\widetilde{w}_{j,r}\}, \{k_{jr}\}, \psi_j(\widetilde{u}_j), \ \delta_j): \text{ one has } k_{jr} \to \infty,$$

$$\widetilde{w}_{j,r}(\alpha_j(k_{jr}^{-1}\widetilde{u}_j)) \to \psi_j(\widetilde{u}_j) \text{ uniformly with derivatives on compact subsets in } \mathbb{R}^{d_j}, \tag{6.24}$$

$$\psi_j : \overline{D}_{\delta_j} \to \psi_j(\overline{D}_{\delta_j}) \subset H_j \text{ is a diffeomorphism.} \tag{6.25}$$

Fix a $R > 0$ such that

$$\widetilde{w}_{j,r}^{-1}(\alpha_j(0)) = \widetilde{w}_{j,r}^{-1}(A_{(j)}, B_{(j)}) \in D_R \subset H_j \text{ for any } j \text{ and } r,$$

which exists by (6.24). Let

$$\Omega_{m,D_R}, \ \widetilde{l}_m = l_m(D_R), \ c_R = c(A, B, D_R)$$

be as in (6.5) (they correspond to the $\varepsilon(x)$- approximations on the ball $D_R \subset G$ by words in (A, B)). Recall that the subset $\Omega_{m,D_R}(A_{(j)}, B_{(j)}) \subset H_j$ contains a $\varepsilon(c\widetilde{l}_m)$- net on $D_R \subset H_j$ (as at the beginning of the subsection). Fix some words

$$\omega_{(j)rm} \in \Omega_{m,D_R} \text{ such that } dist(\omega_{(j)rm}(A_{(j)}, B_{(j)}), \widetilde{w}_{j,r}^{-1}(A_{(j)}, B_{(j)})) < \varepsilon(c_R\widetilde{l}_m), \tag{6.26}$$

as in (6.6). Put

$$w'_{(j)m} = \widetilde{w}_{j,r}\omega_{(j)rm}, \tag{6.27}$$

fixing r large enough, so that there exist constants $K, \Delta > 0$ such that for any $j = 1, \dots, s$, any $m \in \mathbb{N}$ (large enough) and any $\eta \in \cup_l \Omega_l$, see (6.19) (including the trivial word)

$$\psi_{\eta,j,m}(\widetilde{u}_j) = (\eta^{-1}w'_{(j)m})(\alpha_j(k_{jr}^{-1}\widetilde{u}_j)) : \overline{D_{\delta_j}} \to \psi_{\eta,j,m}(\overline{D_{\delta_j}}) \subset H_j \text{ is a diffeomorphism,} \tag{6.28}$$

$$||(\psi'_{\eta,j,m}(x))^{-1}|| < K \text{ for any } x \in \overline{D_{\delta_j}}, \tag{6.29}$$

$$D_\Delta \subset (\psi_{\eta,j,m}(0))^{-1}\psi_{\eta,j,m}(D_{\delta_j}) \subset H_j. \tag{6.30}$$

Let us prove the possibility of the previous choice of r. One has

$$\psi_{\eta,j,m}(\widetilde{u}_j) = (\eta^{-1}\widetilde{w}_{j,r}\omega_{(j)rm})(\alpha_j(k_{jr}^{-1}\widetilde{u}_j)), \eta \in \cup_k\Omega_k, \ \omega_{(j)rm} \in \Omega_{m,D_R}. \tag{6.31}$$

If r is large enough, then

- the mapping $\widetilde{w}_{j,r}(\alpha_j(k_{jr}^{-1}\widetilde{u}_j))$ is uniformly close on $\overline{D_{\delta_j}}$ (with derivatives) to the diffeomorphism $\psi_j(\widetilde{u}_j)\overline{D_{\delta_j}} \to \psi_j(\overline{D_{\delta_j}})$, see (6.24), (6.25),

- the numbers k_{jr} are so large that the mappings

$$\widetilde{u}_j \mapsto w(\alpha_j(k_{jr}^{-1}\widetilde{u}_j)), \ w \in (\cup_k \Omega_k^{-1}) \cup (\cup_k \Omega_{k,D_R}),$$

are uniformly close (with derivatives) on $\overline{D_{\delta_j}}$ to the constant mappings $\widetilde{u}_j \mapsto w(\alpha_j(0))$, since the corresponding mappings $w(\alpha_j(u_j))$ have uniformly bounded derivatives on a neighborhood of 0 independent on w (the $\varepsilon(x)$- approximability with bounded derivatives).

The two last statements together with (6.31) imply that if r is large enough, then statement (6.28) holds true and there exist constants $K, \Delta > 0$ such that statements (6.29) and (6.30) hold true for all m and η.

The construction of the word sequences $\eta_{i,m}$, $i = 1, 2$. Let C be the constant from Proposition 6.1. Let Ω_m, l_m, c be as in (6.19). Fix a sequence of word pairs

$$\eta_{1,m}, \eta_{2,m} \in \Omega_m \text{ such that for any } m \in \mathbb{N} \text{ (large enough)}$$

$$\eta_{i,m}(A,B) \in D_{(C+1)\hat{\delta}_m} \subset G, \ \hat{\delta}_m = \varepsilon(cl_m), \text{ for any } i = 1, 2; \ [\eta_{1,m}, \eta_{2,m}](A,B) \neq 1. \tag{6.32}$$

Recall that the subset $\Omega_m(A,B) \subset G$ contains a $\hat{\delta}_m$- net on $D_1 \subset G$. The words $\eta_{1,m}$, $\eta_{2,m}$ exist by Proposition 6.1 applied to the subset in $\Omega_m(A,B)$ that forms a $\hat{\delta}_m$- net on $D_{C\hat{\delta}_m} \subset G$ (the latter $\hat{\delta}_m$- net exists, whenever m is large enough so that $C\hat{\delta}_m \leq 1$). Let $w'_{(j)m}$ be the word sequences constructed above, see (6.27). Fix a sequence

$$i(j,m) = 0, 1, 2, \text{ put } \eta_{0,m} = 1, \ w_{(j)m} = \eta_{i(j,m),m}^{-1} w'_{(j)m}, \tag{6.33}$$

such that the corresponding commutators w_m from (6.22) be nontrivial. (This sequence exists by (6.32) and Proposition 4.2.) Let us show that the words $w_{(j)m}$ and appropriate pairs $(A_{(j)m}, B_{(j)m}) \in H_j \times H_j$ satisfy (6.20) and (6.21).

To do this, we consider the auxiliary mappings

$$\widetilde{\psi}_{(j)m}(\widetilde{u}_j) = w_{(j)m}(\alpha_j(k_{jr}^{-1}\widetilde{u}_j)) = \psi_{\eta_{i(j,m),m},j,m}(\widetilde{u}_j), \ \widetilde{\psi}_{(j)m}(0) = w_{(j)m}(A_{(j)}, B_{(j)}),$$

which satisfy statements (6.28)-(6.30) by definition. We use the inequality

$$dist(\widetilde{\psi}_{(j)m}(0), 1) < (C+4)\varepsilon(\widetilde{c}l_m), \text{ whenever } m \text{ is large enough,} \tag{6.34}$$

$$\widetilde{c} = \min\{c, c_R\}, \ c \text{ is the constant from (6.19)}, \ c_R = c(A, B, D_R), \text{see (6.5)}.$$

Proof of (6.34). One has

$$\widetilde{\psi}_{(j)m}(0) = (\eta_{i(j,m),m}^{-1} w'_{(j)m})(A_{(j)}, B_{(j)}),$$

$$\eta_{i,m}(A_{(j)}, B_{(j)}) \in D_{(C+1)\hat{\delta}_m} \subset H_j, \ \hat{\delta}_m = \varepsilon(cl_m),$$

by (6.32) and the definition of the metric of G,

$$\eta_{i,m}^{-1}(A_{(j)}, B_{(j)}) \in D_{(C+3)\hat{\delta}_m} \text{ for any } m \text{ large enough and } i = 0, 1, 2, \tag{6.35}$$

by the previous inclusion and since $\hat{\delta}_m \to 0$ and the inversion $H_j \to H_j : g \mapsto g^{-1}$ has unit derivative at 1 (in particular, an inclusion $g \in D_{(C+1)\hat{\delta}_m}$ implies $g^{-1} \in D_{(C+3)\hat{\delta}_m}$ for any m large enough). One has

$$w'_{(j)m}(A_{(j)}, B_{(j)}) \in D_{\varepsilon(c_R\widetilde{l}_m)} \subset D_{\varepsilon(c_R l_m)} \subset H_j, \text{ where } \widetilde{l}_m \text{ is the same, as in (6.5).} \quad (6.36)$$

The first inclusion follows from (6.26), (6.27) and the left-invariance of the metric of H_j. The second one follows from the inequality $\widetilde{l}_m \geq l_m$, see (6.17). Inclusions (6.35) and (6.36) imply (6.34). □

Let us prove (6.20). The image $\widetilde{\psi}_{(j)m}(D_{\delta_j}) \subset H_j$ contains 1, whenever m is large enough, by statements (6.34) and (6.30) and since the right-hand side in (6.34) tends to 0, as $m \to \infty$. Put

$$\widetilde{u}_{jm} = \widetilde{\psi}^{-1}_{(j)m}(1) \in D_\delta, \ (A_{(j)m}, B_{(j)m}) = \alpha_j(k_{jr}^{-1}\widetilde{u}_{jm}). \text{ Then } w_{(j)m}(A_{(j)m}, B_{(j)m}) = 1 \quad (6.37)$$

by definition. One has

$$|w_{(j)m}| \leq |\omega_{(j)rm}| + |\eta_{i(j,m),m}| + |\widetilde{w}_{j,r}| \leq \widetilde{l}_m + l_m + |\widetilde{w}_{j,r}| \leq c''_j l_m, \ c''_j = \sup_m \frac{\widetilde{l}_m + l_m + |\widetilde{w}_{j,r}|}{l_m}. \quad (6.38)$$

The first inequality in (6.38) follows from (6.27) and (6.33). The second one holds true since $|\omega_{(j)rm}| \leq \widetilde{l}_m$ ($\omega_{(j)rm} \in \Omega_{m,D_R}$, see (6.26)) and $|\eta_{i,m}| \leq l_m$ ($\eta_{i,m} \in \Omega_m$). The third inequality (and the finiteness of the constant c''_j) follow from (6.17). Statements (6.37) and (6.38) imply (6.20).

Let us prove (6.21). One has

$$dist((A_{(j)m}, B_{(j)m}), (A_{(j)}, B_{(j)})) = O(\varepsilon(\widetilde{c} l_m)), \text{ as } m \to \infty, \ \widetilde{c} \text{ is the same, as in (6.34),} \quad (6.39)$$

by (6.37) and since $\widetilde{u}_{jm} = O(\varepsilon(\widetilde{c} l_m))$ (by (6.37), (6.34) and (6.29)). There exists a constant $c' > 1$ such that the "O" in (6.39) is less than $c'\varepsilon(\widetilde{c} l_m) < \varepsilon(\hat{c} l_m)$, $\hat{c} = \widetilde{c}(c')^{-1}$, whenever m is large enough (by definition and (1.1)). This implies (6.21) with $c_j = \hat{c}(c''_j)^{-1}$.

The words $w_{(j)m}$ and the pairs $(A_{(j)m}, B_{(j)m}) \in H_j \times H_j$ satisfy statements (6.20) and (6.21). The iterated commutators w_m, see (6.22), of the words $w_{(j)m}$ are nontrivial. This reduces us to the previous Case 1) and finishes the proof of Theorem 1.29 and its Addendum for arbitrary semisimple Lie group G.

6.3 Case of non-semisimple Lie group

Let a Lie group G be not semisimple. Let $H \subset G$ be the maximal normal solvable connected Lie subgroup. The group $G_{ss} = G/H$ is semisimple, hence, it satisfies the statements of Theorem 1.29 and its Addendum (as it was proved in the two previous Subsections). Denote

$$p : G \to G_{ss} \text{ the quotient projection, } A' = p(A), \ B' = p(B).$$

We fix arbitrary left-invariant Riemann metrics on G and G_{ss}. By assumption, the pair $(A', B') \in G_{ss} \times G_{ss}$ is irrational, and the group G_{ss} is $\varepsilon(x)$- approximable (with bounded derivatives) on the unit ball $D_1 \subset G_{ss}$ by words in (A', B'). Let $l_m = l_m(A', B', D_1) \in \mathbb{N}$

be the corresponding length majorant sequence. There exist $c', c'' > 0$, a sequence w'_m of nontrivial words and a sequence $(A'_m, B'_m) \in G_{ss} \times G_{ss}$, $(A'_m, B'_m) \to (A', B')$, as $m \to \infty$, such that

$$|w'_m| \leq l'_m = c'' l_m, \ w'_m(A'_m, B'_m) = 1, \tag{6.40}$$

$$dist((A'_m, B'_m), (A', B')) < \varepsilon(c' l'_m) \tag{6.41}$$

(Theorem 1.29 and its Addendum applied to G_{ss}). For any $m \in \mathbb{N}$ let us choose some representatives $A_m, B_m \in G$ in the classes $A'_m, B'_m \in G_{ss} = G/H$ respectively so that there exists a constant $C > 1$ such that for any $m \in \mathbb{N}$

$$dist(A_m, A) \leq C dist(A'_m, A'), \ dist(B_m, B) \leq C dist(B'_m, B'). \tag{6.42}$$

(We can just choose a cross-section at A to the class $AH \subset G$ and take A_m in this cross-section, and choose B_m analogously.) By definition,

$$w'_m(A_m, B_m) \in H \text{ for all } m.$$

Let s be the commutant chain length of H (see (4.8)). For any m let $w_m = \widetilde{w}'_m$ be the word from Proposition 4.3 applied to $w = w'_m$, $\widetilde{A} = A_m$, $\widetilde{B} = B_m$. The words w_m are nontrivial, $w_m(A_m, B_m) = 1$. By (4.9),

$$|w_m| \leq 4^{s+1} |w'_m| \leq \hat{l}_m = 4^{s+1} l'_m = \widetilde{c}'' l_m, \ \widetilde{c}'' = 4^{s+1} c'',$$

$$dist((A_m, B_m), (A, B)) \leq C dist((A'_m, B'_m), (A', B')) < C\varepsilon(c' l'_m) < \varepsilon(\widetilde{c}' \hat{l}_m), \ \widetilde{c}' = c' C^{-1} 4^{-(s+1)},$$

whenever m is large enough, see (6.41), (6.42) and (1.1). The two last inequalities show that the pair $(A, B) \in G \times G$ is $\varepsilon(x)$- approximable by pairs (A_m, B_m) with relations of lengths at most $\hat{l}_m = \widetilde{c}'' l_m$. The proof of Theorem 1.29 and its Addendum is complete.

7 Approximability of semisimple Lie groups

Here we firstly prove Lemma 1.25. Then we prove Theorem 1.26.

7.1 The weak Solovay-Kitaev inequality. Proof of Lemma 1.25

Let $\mathfrak{g}$ be a semisimple Lie algebra, $x \in \mathfrak{g}$ be a regular element, $\mathfrak{h} \subset \mathfrak{g}$ be the corresponding Cartan subalgebra, $\Delta_{\mathfrak{g}} \subset \mathfrak{h}^*_{\mathbb{C}}$ be the root system of $\mathfrak{h}_{\mathbb{C}}$ in $\mathfrak{g}_{\mathbb{C}}$, see Subsection 1.9. Let $\mathfrak{g}_\alpha \subset \mathfrak{g}_{\mathbb{C}}$ be the complex root eigenlines, $E \subset \mathfrak{g}$ be the complementary subspace to $\mathfrak{h}$ from (1.24):

$$E_{\mathbb{C}} = \oplus_{\alpha \in \Delta_{\mathfrak{g}}} \mathfrak{g}_\alpha; \ \mathfrak{g} = \mathfrak{h} \oplus E, \ x \in \mathfrak{h}, \ \mathrm{ad}_x(E) \subset E.$$

Proposition 7.1 *In the above assumptions the linear operator* $\mathrm{ad}_x : E \to E$ *is an automorphism.*

The proposition follows from definition: the values of the roots at x do not vanish by regularity.

As it is shown below, Lemma 1.25 is implied by the following

Lemma 7.2 *Let $\mathfrak{g}$, $\mathfrak{h}$, E be as above, G be a Lie group with algebra $\mathfrak{g}$. There exists a $g \in G$ such that*

$$\mathfrak{g} = E + Ad_g E.$$

Proof For each $\alpha \in \Delta_{\mathfrak{g}}$ consider the corresponding elements $e_{\pm\alpha} \in \mathfrak{g}_{\pm\alpha}$, $h_\alpha \in \mathfrak{h}_{\mathbb{C}}$, see Subsection 1.9 or [21], p. 159:

$$[e_\alpha, e_{-\alpha}] = h_\alpha. \tag{7.1}$$

Let $S = \{\alpha_1, \dots, \alpha_r\} \subset \Delta_{\mathfrak{g}}$ be a basis of roots, $S_{\mathbb{R}} \subset S$ be the subset of real roots. Denote $S'' \subset S \setminus S_{\mathbb{R}}$ the maximal subset of nonreal basic roots that contains no conjugated root pair, see (1.21):

$$\alpha \neq \widetilde{\beta} \text{ for any } \alpha, \beta \in S''. \text{ Put } S' = S_{\mathbb{R}} \cup S'', \tag{7.2}$$

$V_\alpha = \{\tau h_\alpha \mid \tau \in \mathbb{R}\}$ for any $\alpha \in S_{\mathbb{R}}$, $V_\alpha = \{\tau h_\alpha + \overline{\tau}\overline{h}_\alpha \mid \tau \in \mathbb{C}\}$ for any $\alpha \in S''$. One has

$$V_\alpha \subset \mathfrak{h}, \text{ by (1.22), } \mathfrak{h} = \sum_{\alpha \in S'} V_\alpha.$$

Indeed, the latter sum is the real part of the complex space spanned by the vectors $h_\alpha, h_{\widetilde{a}} = \overline{h}_\alpha \in \mathfrak{h}_{\mathbb{C}}$ with $\alpha \in S'$, $\{\alpha, \widetilde{\alpha} \mid \alpha \in S'\} \supset S$ by definition. Thus, the collection of the previous vectors contains $\{h_\alpha\}_{\alpha \in S}$, which is a basis in $\mathfrak{h}_{\mathbb{C}}$, see (1.20).

Denote $\pi_h : \mathfrak{g} \to \mathfrak{h}$ the projection along E. Consider the following family of group elements depending on a complex parameter:

$$g(t) = \exp v(t) \in G, \ v(t) = \sum_{\alpha \in S'} (te_\alpha + \bar{t}\overline{e}_\alpha), \ t \in \mathbb{C}.$$

(It follows from definition that $v(t)$ is a real vector in $\mathfrak{g}_{\mathbb{C}}$, thus, $v(t) \in \mathfrak{g}$ and $g(t) \in G$.) We show that for any t small enough and any $\alpha \in S'$ the image $\pi_h(Ad_{g(t)}E)$ contains a subspace $\hat{V}_\alpha \subset \mathfrak{h}$ of the same dimension, as V_α that is arbitrarily close to V_α (as a point of the corresponding grassmanian of subspaces in $\mathfrak{h}$). The subspaces $\hat{V}_\alpha$ generate $\mathfrak{h}$, as do V_α, whenever t is small enough. Hence, $Ad_{g(t)}E$ and E generate $\mathfrak{g}$. This will prove the lemma.

For the proof of the previous statement on $\hat{V}_\alpha$ for any $\alpha \in S'$ let us consider the following auxiliary vector family in $\mathfrak{g}$:

$$v_\alpha(\tau) = \tau e_{-\alpha}, \ \tau \in \mathbb{R}, \text{ if } \alpha \in S_{\mathbb{R}}; \ v_\alpha(\tau) = \tau e_{-\alpha} + \overline{\tau}\overline{e}_{-\alpha}, \ \tau \in \mathbb{C}, \text{ if } \alpha \in S''.$$

By definition, the vectors $v_\alpha(\tau)$ thus constructed are real ($e_{\pm\alpha}$ are chosen real for a real α, see 1.9), hence lie in $\mathfrak{g}$. One has

$$\pi_h(Ad_{g(t)}v_\alpha(\tau)) = (t + \bar{t})\tau h_\alpha + O(\tau t^2), \text{ if } \alpha \in S_{\mathbb{R}}; \tag{7.3}$$

$$\pi_h(Ad_{g(t)}v_\alpha(\tau)) = t\tau h_\alpha + \bar{t}\overline{\tau}\overline{h}_\alpha + O(\tau t^2), \text{ if } \alpha \in S''. \tag{7.4}$$

Indeed, in the case, when $\alpha \in S_{\mathbb{R}}$, the first order term (in t) of $Ad_{g(t)}v_\alpha(\tau)$ is

$$[v(t), v_\alpha(\tau)] = (t + \bar{t})\tau h_\alpha + \tau \sum_{\beta \in S' \setminus \alpha} (t[e_\beta, e_{-\alpha}] + \bar{t}[\overline{e}_\beta, e_{-\alpha}]), \text{ by (7.1).}$$

The latter commutators (which are complex conjugate) lie in $E_{\mathbb{C}}$. Indeed, by (1.18), the commutator $[e_\beta, e_{-\alpha}]$ either is zero, or lies in $\mathfrak{g}_{\beta-\alpha} \subset E_{\mathbb{C}}$ ($\beta - \alpha \neq 0$ by definition). This implies (7.3).

Analogously, in the case, when $\alpha \in S''$, the first order term (in t) of $Ad_{g(t)}v_\alpha(\tau)$ equals $t\tau h_\alpha + \bar{t}\bar{\tau}\overline{h}_\alpha$ plus a complex linear combination of the commutators $[\overline{e}_\alpha, e_{-\alpha}]$, $[e_\beta, e_{-\alpha}]$, $[\overline{e}_\beta, e_{-\alpha}]$ (with $\beta \neq \alpha$) and their complex conjugates. The latter commutators lie in $E_{\mathbb{C}}$. Indeed, by (1.18) and (1.22), each of them is either zero, or lies in $\mathfrak{g}_\gamma \subset E_{\mathbb{C}}$, $\gamma = \widetilde{\alpha}-\alpha, \beta-\alpha, \widetilde{\beta}-\alpha$ respectively (if $\gamma \neq 0$), or lies in $\mathfrak{h}$, if $\gamma = 0$. But the latter case is impossible: $\widetilde{\alpha} - \alpha \neq 0$ ($\alpha \in S''$ is not a real root), $\beta - \alpha \neq 0$ (by definition), $\widetilde{\beta} - \alpha \neq 0$. Indeed, $\beta \in S' = S_{\mathbb{R}} \cup S''$. If $\beta \in S_{\mathbb{R}}$, then $\widetilde{\beta} - \alpha \neq 0$, since α is not a real root. If $\beta \in S''$, then $\widetilde{\beta} - \alpha \neq 0$, by the definition of the set S'', see (7.2). This together with the previous discussion implies (7.4).

Formulas (7.3) and (7.4) imply that the image under $\pi_h \circ Ad_{g(t)}$ of the vector space $W_\alpha = \{v_\alpha(\tau)\} \subset \mathfrak{g}$ contains a subspace $\hat{V}_\alpha$ of the same dimension, as V_α, approaching the latter, as $t \to 0$. This together with the previous discussion proves the lemma. □

Let us prove Lemma 1.25. To do this, fix a $g \in G$ as in Lemma 7.2. Denote

$$E^1 = E, \ E^2 = Ad_g E. \text{ Fix a subspace } T \subset E^2 \text{ such that } \mathfrak{g} = E^1 \oplus T.$$

We write each $z \in \mathfrak{g}$ as

$$z = z_1 + z_2, \ z_1 \in E^1, \ z_2 \in T.$$

There exists a $K > 0$ such that

$$\max_j |z_j| \leq K|z| \text{ for any } z \in \mathfrak{g}. \tag{7.5}$$

This follows from definition. Let us choose a regular element

$$\widetilde{x}_1 \in \mathfrak{h}; \text{ then the operator } \operatorname{ad}_{\widetilde{x}_1} : E^1 \to E^1 \text{ is nondegenerate. Put } \widetilde{x}_2 = Ad_g \widetilde{x}_1.$$

Then the space E^2 is invariant under the action of the operator $\operatorname{ad}_{\widetilde{x}_2}$, and the latter being restricted to E^2 is nondegenerate. One has

$$z_j = [\widetilde{x}_j, \widetilde{y}_j], \ \widetilde{y}_j = (\operatorname{ad}_{\widetilde{x}_j}|_{E^j})^{-1} z_j, \ |\widetilde{y}_j| \leq K_j |z_j|, \ K_j = ||(\operatorname{ad}_{\widetilde{x}_j}|_{E^j})^{-1}||.$$

The elements

$$x_j = \widetilde{x}_j \sqrt{\frac{|\widetilde{y}_j|}{|\widetilde{x}_j|}}, \ y_j = \widetilde{y}_j \sqrt{\frac{|\widetilde{x}_j|}{|\widetilde{y}_j|}}$$

satisfy statement (1.12) with $c = \max_j \sqrt{KK_j|\widetilde{x}_j|}$, which follows from definition and (7.5). This proves Lemma 1.25.

7.2 Approximations with weak Solovay-Kitaev property. Proof of Theorem 1.26

Firstly we prove Theorem 1.26. Its proof given below is similar to that of Theorem 1.21 in [6, 15, 17]. The proof of the boundedness of derivatives in Theorem 1.21 (see Remark 1.22) will be given at the end of the subsection (it is a small modification of the proof of Theorem 1.26).

For any collection Ω of either words or elements of a Lie group denote

$$\Omega'' = \{[x_1, y_1][x_2, y_2] \mid x_i, y_i \in \Omega\}.$$

Let G be a Lie group whose Lie algebra satisfies the weak Solovay-Kitaev inequality (1.12), c be the constant from (1.12). The group G is equipped with a left-invariant Riemann metric. Let $(A,B)\in G\times G$ be an arbitrary irrational pair. For the proof of Theorem 1.26 it suffices to show that the group G is $\varepsilon(x)$- approximable on the unit ball $D_1\subset G$ by words in (A,B) with bounded derivatives, where the function $\varepsilon(x)$ and the length majorants $l_m=l_m(A,B,D_1)$ are the same as in (1.6) and (1.7). We construct the corresponding word collections $\Omega_m=\Omega_{m,D_1}$ below by induction in m so that each word from Ω_{m+1} is a product of a word from Ω_m and some (iterated) commutators of words from $\cup_{k\leq m}\Omega_k$. To show that $\Omega_m(A,B)$ contains a net on D_1 with appropriate radius, we use the following lemma. It implies that if $\Omega\subset G$ is a δ-net on D_1, then the product $\Omega\Omega''$ contains a $c''\delta^{\frac{3}{2}}$- net on D_1, $c''>0$ is a constant depending only on G, the constant δ is arbitrary (small enough).

Lemma 7.3 *Let G be a Lie group whose Lie algebra satisfies the weak Solovay-Kitaev inequality (1.12). There exist constants $c',c''>0$ such that for any $\delta>0$ small enough and any δ- net Ω on $D_{c'\sqrt{\delta}}\subset G$ the set Ω'' contains a $c''\delta^{\frac{3}{2}}$- net on D_δ. (The constant c' may be chosen arbitrarily close to the corresponding constant c from inequality (1.12), whenever δ is small enough.)*

Proof Fix arbitrary $c'>c$. We consider that δ is so small that the ball $D_\delta\subset G$ is covered by the exponential mapping $\mathfrak{g}\to G$. Fix a δ- net Ω on $D_{c'\sqrt{\delta}}\subset G$. It is sufficient to show that there exists a $c''>0$ (depending only on G, neigher on δ, nor on Ω) such that for any $z\in D_\delta\subset G$ there exist $x_1,y_1,x_2,y_2\in\Omega$ such that

$$dist(z,[x_1,y_1][x_2,y_2])<c''\delta^{\frac{3}{2}}. \tag{7.6}$$

Fix a $z\in D_\delta$. We use the asymptotic equality

$$dist(\exp\hat{v},1)=|\hat{v}|(1+o(1)),\ \text{as}\ \hat{v}\to 0;\ \hat{v}\in\mathfrak{g}, \tag{7.7}$$

which holds true, since the exponential mapping has unit derivative at 1. We write

$$z=\exp(v),\ v\in\mathfrak{g},\ |v|=dist(z,1)(1+o(1))\leq\delta(1+o(1)),\ \text{as}\ \delta\to 0, \tag{7.8}$$

$$v=[u_1,v_1]+[u_2,v_2],\ u_i,v_i\in\mathfrak{g},\ |u_i|=|v_i|\leq c\sqrt{|v|}\leq c\sqrt{\delta}(1+o(1))<c'\sqrt{\delta}, \tag{7.9}$$

whenever δ is small enough (dependently on G and the choice of c'), by (1.12) and (7.8). Put

$$\widetilde{x}_i=\exp u_i,\ \widetilde{y}_i=\exp v_i.\ \text{One has}\ \widetilde{x}_i,\widetilde{y}_i\in D_{c'\sqrt{\delta}}, \tag{7.10}$$

by (7.9) and the left invariance of the metric. Fix

$$x_i,y_i\in\Omega:\ \text{some}\ \delta-\ \text{approximants of}\ \widetilde{x}_i\ \text{and}\ \widetilde{y}_i\ \text{respectively.} \tag{7.11}$$

These are the x_i and the y_i we are looking for. Indeed, first let us show that

$$dist(z,[\widetilde{x}_1,\widetilde{y}_1][\widetilde{x}_2,\widetilde{y}_2])=O(\delta^{\frac{3}{2}}),\ \text{as}\ \delta\to 0. \tag{7.12}$$

Proof of (7.12). We use the following asymptotic relations between products and commutators in Lie group and sums and brackets in its Lie algebra:

$$dist(\exp(a+b),\exp a\exp b)=O(|a||b|),\ \text{as}\ a,b\to 0, \tag{7.13}$$

$$dist(\exp([u,v]),[\exp u,\exp v]) = O(|u|^3+|v|^3), \text{ as } u,v\to 0. \tag{7.14}$$

Applying (7.13) to $a=[u_1,v_1]$ and $b=[u_2,v_2]$ (then $\exp(a+b)=z$, see (7.8) and (7.9)) and writing $[u_i,v_i]=O(|u_i||v_i|)=O(\delta)$, see (7.9), yields

$$dist(z,\exp([u_1,v_1])\exp([u_2,v_2])) = O(\delta^2).$$

Applying (7.14) to $u=u_i$, $v=v_i$ and writing $|u_i|=|v_i|=O(\sqrt{\delta})$, see (7.9), yields

$$dist(\exp([u_i,v_i]),[\widetilde{x}_i,\widetilde{y}_i]) = O(\delta^{\frac{3}{2}}).$$

Substituting the latter equality to the previous one yields (7.12). □

One has

$$dist(z,[x_1,y_1][x_2,y_2]) = O(\delta^{\frac{3}{2}}), \text{ as } \delta\to 0. \tag{7.15}$$

Indeed, for any $i=1,2$ one has

$$dist([x_i,y_i],[\widetilde{x}_i,\widetilde{y}_i]) \leq dist([x_i,y_i],[x_i,\widetilde{y}_i]) + dist([x_i,\widetilde{y}_i],[\widetilde{x}_i,\widetilde{y}_i])$$

$$= O(dist(y_i,\widetilde{y}_i)dist(x_i,1)) + O(dist(x_i,\widetilde{x}_i)dist(\widetilde{y}_i,1)) = O(\delta^{\frac{3}{2}}),$$

by (7.11) and (7.10). This together with (7.12) implies (7.15). There exists a constant $c''>0$ (depending only on G) such that the "O" in (7.15) is less than $c''\delta^{\frac{3}{2}}$, whenever δ is small enough (independently on Ω). This proves (7.6) and Lemma 7.3. □

Let c',c'' be the constants from Lemma 7.3. Without loss of generality everywhere below we consider that

$$c',c''>1. \tag{7.16}$$

Fix a $0<\delta<1$ small enough that satisfies the statements of Lemma 7.3 and the inequality

$$c''\delta^{\frac{3}{2}} < \delta. \tag{7.17}$$

Consider the positive number sequence $\delta_m>0$ defined recurrently as follows:

$$\delta_1=\delta, \ \delta_{m+1}=c''\delta_m^{\frac{3}{2}}. \tag{7.18}$$

The sequence δ_m decreases to 0 superexponentially, as $m\to\infty$, by (7.17). Fix a finite word collection Ω so that

$$\Omega(A,B) \text{ is a } \delta- \text{ net on } D_1\subset G.$$

We define sequences Ω_m, $\widetilde{\Omega}_m$ of word collections by induction as follows:

$$\Omega_1=\Omega, \ \widetilde{\Omega}_1=\{w\in\Omega_1 \mid w(A,B)\in D_{2c'\sqrt{\delta_1}}\subset G\}, \tag{7.19}$$

$$\Omega_2=\Omega_1\widetilde{\Omega}_1'', \widetilde{\Omega}_2=\{w\in\Omega_2 \mid w(A,B)\in D_{2c'\sqrt{\delta_2}}\subset G\}, \tag{7.20}$$

$$\Omega_{m+1}=\Omega_m\widetilde{\Omega}_m'', \ \widetilde{\Omega}_{m+1}=\widetilde{\Omega}_{m-1}''\widetilde{\Omega}_m'' \text{ for any } m\geq 2. \tag{7.21}$$

We show that the sequence of collections Ω_m, the set $K=D_1\subset G$ and the numbers

$$l_m=9^{m-1}l_1, \ l_1=\max_{w\in\Omega}|w|, \tag{7.22}$$

satisfy the statements of Definition 1.11 (the $\varepsilon(x)$- approximability on D_1 with bounded derivatives), whenever δ is small enough. To do this, it suffices to show that

$$|w| \leq l_m \text{ for any } w \in \Omega_m \text{ and any } m \in \mathbb{N}, \tag{7.23}$$

$$\text{the set } \Omega_m(A,B) \text{ contains a } \delta_m - \text{ net on } D_1 \text{ for all } m \in \mathbb{N}, \tag{7.24}$$

$$\text{there exists a } \hat{c} > 0 \text{ (depending only on } \delta \text{ and } \Omega) \text{ such that } \delta_m < \varepsilon(\hat{c}l_m) \text{ for all } m \in \mathbb{N}, \tag{7.25}$$

$$\text{the subset } \cup_m \Omega_m(A,B) \subset G \text{ is bounded}, \tag{7.26}$$

- there exists a neighborhood $U \subset G \times G$ of (A,B) where the mappings

$$w : (a,b) \mapsto w(a,b),\ U \to G,\ w \in \cup_m \Omega_m, \text{ have uniformly bounded derivatives.} \tag{7.27}$$

Statements (7.24) and (7.25) imply that the set $\Omega_m(A,B)$ contains a $\varepsilon(\hat{c}l_m)$- net on D_1. This together with (7.23), (7.26) and (7.27) proves the $\varepsilon(x)$- approximability of G on D_1 by words in (A,B) with bounded derivatives. This will prove Theorem 1.26.

Statements (7.23)-(7.27) are proved below. Statement (7.23) will be proved for arbitrary δ, while statements (7.24)-(7.27) will be proved for any δ small enough (as it will be specified at the beginning of the proof of each one of these statements).

Proof of (7.23). We prove (7.23) and the auxiliary inequality

$$|w| \leq l_m \text{ for any } w \in \widetilde{\Omega}_m \tag{7.28}$$

by induction in m.

Induction base. For $m = 1$ inequality (7.23) follows from definition, and inequality (7.28) is its particular case, see (7.19).

Induction step. Let (7.23), (7.28) hold true for all $m \leq j$. Let us prove them for $m = j+1$. The induction hypothesis implies that

$$|w| \leq 8l_k \text{ for any } w \in \Omega''_k \cup \widetilde{\Omega}''_k,\ k \leq j. \tag{7.29}$$

Let us firstly prove (7.23). For any $w \in \Omega_{j+1}$ one has

$$w = w_1 w_2,\ w_1 \in \Omega_j,\ w_2 \in \widetilde{\Omega}''_j, \text{ see (7.20) and (7.21)}, |w_1| \leq l_j,\ |w_2| \leq 8l_j$$

(the induction hypothesis and (7.29)), $l_{j+1} = 9l_j$. Therefore, $|w| \leq l_{j+1}$. Inequality (7.23) for $m = j+1$ is proved.

Now let us prove inequality (7.28). Let $w \in \widetilde{\Omega}_{j+1}$. If $j+1 = 2$, then $\widetilde{\Omega}_{j+1} \subset \Omega_{j+1}$, see (7.20), and inequality (7.23) (already proved) implies (7.28). Let now $j+1 \geq 3$. Then

$$w = w_1 w_2,\ w_1 \in \widetilde{\Omega}''_{j-1},\ w_2 \in \widetilde{\Omega}''_j,\ |w_1| \leq 8l_{j-1},\ |w_2| \leq 8l_j,$$

by (7.29). This together with the inequality

$$8l_{j-1} + 8l_j = (1 - \frac{1}{9})l_j + 8l_j < 9l_j = l_{j+1}$$

proves (7.28). The induction step is over. Inequalities (7.23) and (7.28) are proved. □

Proof of (7.24). One has

$$2c'\sqrt{\delta} < 1, \ \delta_m < c'\sqrt{\delta}_m, \ \delta_{m-2} > c'\sqrt{\delta}_m \text{ for any } m \in \mathbb{N}, \tag{7.30}$$

whenever δ is small enough. We prove (7.24) for those δ that satisfy statements of Lemma 7.3 and (7.30). We use and prove simultaneously (by induction in m) that

$$\text{the set } \widetilde{\Omega}''_m(A,B) \text{ contains a } \delta_{m+1} - \text{ net on } D_{\delta_m} \text{ for any } m. \tag{7.31}$$

Induction base: m=1,2. The set $\Omega_1(A,B)$ is a δ_1- net on D_1 by definition. This proves (7.24) for $m=1$. Let us prove (7.31) for $m=1$. The set $\widetilde{\Omega}_1(A,B)$ consists of those elements of the previous δ_1- net that lie in the ball $D_{2c'\sqrt{\delta_1}}$ (by definition). Therefore, $\widetilde{\Omega}_1(A,B)$ contains a δ_1- net on $D_{2c'\sqrt{\delta_1}-\delta_1} \supset D_{c'\sqrt{\delta_1}}$ (the latter inclusion follows from the middle inequality in (7.30)). This together with Lemma 7.3 (applied to $\widetilde{\Omega}_1(A,B)$ and δ_1) implies that the set $\widetilde{\Omega}''_1(A,B)$ contains a $\delta_2 = c''\delta_1^{\frac{3}{2}}$- net on D_{δ_1} and proves (7.31) for $m=1$.

The set $\Omega_2(A,B) = \Omega_1(A,B)\widetilde{\Omega}''_1(A,B)$ contains a δ_2- net on D_1, since $\Omega_1(A,B)$ is a δ_1- net on D_1, $\widetilde{\Omega}''_1(A,B)$ contains a δ_2- net on D_{δ_1} (as was shown above) and by Proposition 1.9. This proves (7.24) for $m=2$. The set $\widetilde{\Omega}''_2(A,B)$ contains a δ_3- net on D_{δ_2}, analogously to the previous similar statement on $\widetilde{\Omega}''_1(A,B)$ proved above. This proves statement (7.31) for $m=2$.

Induction step: $m \geq 3$. Let statements (7.24), (7.31) hold true for the smaller indices (the induction hypothesis). Let us prove them for the given m. The collection

$$\Omega_m(A,B) = \Omega_{m-1}(A,B)\widetilde{\Omega}''_{m-1}(A,B)$$

contains a δ_m- net on D_1, which follows immediately from the induction hypothesis for $\Omega_{m-1}(A,B)$, $\widetilde{\Omega}''_{m-1}(A,B)$, as in the above proof of the similar statement for $\Omega_2(A,B)$. This proves (7.24). The set $\widetilde{\Omega}_m(A,B) = \widetilde{\Omega}''_{m-2}(A,B)\widetilde{\Omega}''_{m-1}(A,B)$ contains a δ_m- net on $D_{\delta_{m-2}}$, since $\widetilde{\Omega}''_{m-2}(A,B)$ contains a δ_{m-1}- net on $D_{\delta_{m-2}}$, $\widetilde{\Omega}''_{m-1}(A,B)$ contains a δ_m- net on $D_{\delta_{m-1}}$ (the induction hypothesis), and by Proposition 1.9. One has $\delta_{m-2} > c'\sqrt{\delta_m}$, see (7.30). Therefore, the set $\widetilde{\Omega}''_m(A,B)$ contains a $\delta_{m+1} = c''\delta_m^{\frac{3}{2}}$- net on D_{δ_m} (Lemma 7.3 applied to δ_m and to the previous δ_m- net on $D_{\delta_{m-2}} \supset D_{c'\sqrt{\delta_m}}$). This proves (7.31). The induction step is over. Statements (7.24) and (7.31) are proved. □

Proof of (7.25). We prove (7.25) for any δ satisfying the inequality

$(c'')^2\delta < 1$. Denote $q = -\ln((c'')^2\delta) > 0$. More precisely, we show that

$$\delta_m < \varepsilon(\hat{c}l_m) = e^{-(\hat{c}l_m)^\kappa} \text{ for any } m \in \mathbb{N}, \text{ where } \kappa = \frac{\ln 1.5}{\ln 9}, \ \hat{c} = \frac{q^{\frac{1}{\kappa}}}{l_1}. \tag{7.32}$$

Indeed, the sequence $\ln\delta_m$ is the orbit of $\ln\delta_1 = \ln\delta$ under the affine mapping

$L : x \mapsto \frac{3}{2}x + \ln c''$, which has a fixed point $x_0 = -2\ln c'' < 0$, see (7.16). Hence,

$$\ln\delta_m = x_0 + (\frac{3}{2})^{m-1}(\ln\delta - x_0) = x_0 - (\frac{3}{2})^{m-1}q, \tag{7.33}$$

since $\ln\delta - x_0 = \ln\delta + 2\ln c'' = -q$ by definition. One has

$$\ln\delta_m < -(\frac{3}{2})^{m-1}q = -9^{\kappa(m-1)}q = -(\frac{l_m}{l_1})^{\kappa}q,$$

by definition, (7.33) and the inequality $x_0 < 0$. Taking the exponent of the previous inequality yields (7.32). This proves (7.25). □

Proof of (7.26) and (7.27). For the proof of (7.26) and (7.27) we consider a complex neighborhood $\hat{G}_{\mathbb{C}} \supset G$ of G equipped with the natural structure of (local) complex Lie group. The left-invariant Riemann metric on G extends up to a left-invariant Hermitian metric on $\hat{G}_{\mathbb{C}}$. By definition, one has

$$\Omega_m = \Omega_1\Omega'_{m-1},\ \Omega'_k = \widetilde{\Omega}''_1 \dots \widetilde{\Omega}''_k. \tag{7.34}$$

The collections Ω_m, Ω'_m depend on the choice of δ. Below we define a continuous function

$$\tau = \tau(\delta),\ \tau(0) = 0,$$

see (7.36), for small δ and show that if δ is small enough, then there exists a complex neighborhood $U' \subset \hat{G}_{\mathbb{C}} \times \hat{G}_{\mathbb{C}}$ of (A, B) such that each mapping $w : (a, b) \mapsto w(a, b)$, $w \in \cup_m\Omega'_m$, extends up to a holomorphic mapping

$$w : U' \to D_{\tau(\delta)} \subset \hat{G}_{\mathbb{C}}. \tag{7.35}$$

We take δ so small that the closed ball $\overline{D}_{\tau(\delta)} \subset \hat{G}_{\mathbb{C}}$ be covered by a holomorphic chart of $\hat{G}_{\mathbb{C}}$. The mappings (7.35) are uniformly bounded holomorphic mappings $U' \to \mathbb{C}^n$ in the latter chart. Take arbitrary smaller (real) neighborhood $U \subset G \times G$ of (A, B) such that $\overline{U} \subset U'$. The mappings (7.35) are uniformly bounded with derivatives on U by the previous statement and the Cauchy estimate. This together with (7.34) implies (7.26) and (7.27).

Thus, the previous discussion proves (7.26) and (7.27) modulo statement (7.35). In the proof of (7.35) we use the following obvious

Proposition 7.4 *Let G be a (local) Lie group (equipped with a Riemann metric). There exist constants $\widetilde{c}, \sigma > 0$ such that for any $0 < \tau < \sigma$ and any $x, y \in D_\tau \subset G$ one has*

$$dist([x, y], 1) \leq \widetilde{c}dist(x, 1)dist(y, 1),\ dist(xy, 1) \leq 3\tau,\ dist(xyx^{-1}, 1) \leq 4\tau.$$

Proof of (7.35). Let c' be the constant from (7.19), $\widetilde{c}$, σ be the constants from Proposition 7.4 applied to $\hat{G}_{\mathbb{C}}$. Let us consider the number sequence $\tau_m > 0$ defined recurrently as follows:

$$\tau_1 = \tau_2 = 4c'\sqrt{\delta},\ \tau_j = 4\widetilde{c}\max\{\tau_{j-2}^2, \tau_{j-1}^2\}. \text{ Put } \tau = \tau(\delta) = \sum_{j=1}^{\infty}\tau_j. \tag{7.36}$$

Recall that $\hat{G}_{\mathbb{C}}$ is a local complex Lie group: an open subset of a complex manifold M with a (holomorphic) multiplication operation $\hat{G}_{\mathbb{C}} \times \hat{G}_{\mathbb{C}} \to M$. In general, $\hat{G}_{\mathbb{C}}$ is not necessarily a complete metric space (in its left-invariant Hermitian metric): the distance of its boundary $\partial\hat{G}_{\mathbb{C}}$ to the unity may be finite.

Let δ be small enough so that τ_j decrease to 0 (then they decrease superexponentially and their sum $\tau(\delta)$ is a well-defined continuous function in small δ, $\tau(0)=0$), and

$$\tau(\delta) < \min\{\sigma, \frac{1}{6}dist(1, \partial\hat{G}_{\mathbb{C}})\}. \tag{7.37}$$

The latter inequality implies that the closed balls

$$\overline{D}_{\tau_j} \subset \overline{D}_{5\tau} \subset \hat{G}_{\mathbb{C}} \text{ are compact subsets in } \hat{G}_{\mathbb{C}}. \tag{7.38}$$

Take a complex neighborhood $U' \subset \hat{G}_{\mathbb{C}} \times \hat{G}_{\mathbb{C}}$ of (A, B) such that each word $w \in \widetilde{\Omega}_1 \cup \widetilde{\Omega}_2$ extends up to a holomorphic mapping

$$w : U' \to \hat{G}_{\mathbb{C}},\ w(U') \subset D_{\tau_1} = D_{\tau_2} \subset \hat{G}_{\mathbb{C}} \text{ for any } w \in \widetilde{\Omega}_1 \cup \widetilde{\Omega}_2. \tag{7.39}$$

The existence of U' follows from (7.19), (7.20) and (7.36). We prove (7.35) for the above U' and τ. To do this, we firstly show that for any $m \in \mathbb{N}$ each $w \in \widetilde{\Omega}_m \cup \widetilde{\Omega}''_m$ defines a holomorphic mapping $w : U' \to \hat{G}_{\mathbb{C}}$ and

$$\widetilde{\Omega}_m(U') \subset D_{\tau_m} \subset \hat{G}_{\mathbb{C}},\ \widetilde{\Omega}''_{m-2}(U') \subset D_{2\widetilde{c}\tau^2_{m-2}} \subset \hat{G}_{\mathbb{C}}. \tag{7.40}$$

Proof of (7.40). We prove (7.40) by induction. In the proof we use the inequality

$$2\widetilde{c}\tau^2_{m-2} + 2\widetilde{c}\tau^2_{m-1} \leq \tau_m < \tau. \tag{7.41}$$

Indeed, $4\widetilde{c}\tau^2_{m-2}, 4\widetilde{c}\tau^2_{m-1} \leq \tau_m < \tau$ (by (7.36)), which implies (7.41).

Induction base. For $m = 1, 2$ statement (7.40) follows from (7.39) (the right inclusion in (7.40) is void, since $m - 2 \leq 0$).

Induction step: $m \geq 3$. Let the induction hypothesis hold true: (7.40) holds for smaller indices. One has

$$\widetilde{\Omega}_m = \widetilde{\Omega}''_{m-2}\widetilde{\Omega}''_{m-1}, \tag{7.42}$$

by definition. Let us show that for any $j \leq m-1$ each $w \in \widetilde{\Omega}''_j$ defines a holomorphic mapping $U' \to \hat{G}_{\mathbb{C}}$,

$$\widetilde{\Omega}''_j(U') \subset D_{2\widetilde{c}\tau^2_j}. \tag{7.43}$$

Proof of (7.43). Each $w \in \widetilde{\Omega}_j$ defines a holomorphic mapping $U' \to \hat{G}_{\mathbb{C}}$ and

$$\widetilde{\Omega}_j(U') \subset D_{\tau_j} \tag{7.44}$$

(the induction hypothesis of (7.40)). For any $x_1, y_1, x_2, y_2 \in \Omega_j$ the commutator product $[x_1, y_1][x_2, y_2]$ defines a holomorphic mapping $U' \to \hat{G}_{\mathbb{C}}$. Indeed, for any $a, b \in D_{\tau_j} \subset \hat{G}_{\mathbb{C}}$ the products $ab, aba^{-1} \in \hat{G}_{\mathbb{C}}$ are well-defined and

$$ab, aba^{-1} \in D_{4\tau_j} \subset D_{4\tau} \subset \hat{G}_{\mathbb{C}},\ [a, b] \in D_{\widetilde{c}\tau^2_j} \subset D_\tau \subset \hat{G}_{\mathbb{C}} :$$

the former inclusion follows from Proposition 7.4 (with G replaced by $\hat{G}_{\mathbb{C}}$, τ replaced by τ_j); the latter inclusion follows from the same Proposition and (7.41). This implies that each commutator $[x_i, y_i]$ defines a holomorphic mapping $[x_i, y_i] : U' \to D_{\widetilde{c}\tau^2_j}$. One has $\overline{D}_{2\widetilde{c}\tau^2_j} \subset$

$D_\tau \subset \hat{G}_{\mathbb{C}}$ by (7.41). Hence, the previous commutator product defines a holomorphic mapping $U' \to D_{2\widetilde{c}\tau_j^2}$. This proves (7.43). $\square$

Inclusion (7.43) applied to $j = m-2, m-1$ together with (7.41) and (7.42) imply the first inclusion in (7.40). The second one follows from inclusion (7.43) applied to $j = m-2$. Statement (7.40) is proved. $\square$

For any $m \in \mathbb{N}$ denote

$$T_m = \sum_{j=1}^{m} \tau_j.$$

Now we prove (by induction) that each word $w \in \Omega'_m$ defines a holomorphic mapping $w : U' \to \hat{G}_{\mathbb{C}}$ and

$$\Omega'_m(U') \subset D_{T_m} \text{ for any } m \in \mathbb{N}. \tag{7.45}$$

Proof Induction base: $m = 1$. Each word $w \in \widetilde{\Omega}''_1 = \Omega'_1$ defines a holomorphic mapping $w : U' \to D_{2\widetilde{c}\tau_1^2}$, by the second inclusion in (7.40). One has $2\widetilde{c}\tau_1^2 < \tau_2 = \tau_1$ by definition and (7.41). This shows that $w(U') \subset D_{\tau_1} = D_{T_1}$ and proves (7.45).

Induction step: $m \geq 2$. Let us prove (7.45) assuming it is proved for smaller indices. By definition,

$$\Omega'_m = \Omega'_{m-1}\widetilde{\Omega}''_m. \tag{7.46}$$

Each $w \in \Omega'_{m-1} \cup \widetilde{\Omega}''_m$ defines a holomorphic mapping $U' \to \hat{G}_{\mathbb{C}}$,

$$\Omega'_{m-1}(U') \subset D_{T_{m-1}},\ \widetilde{\Omega}''_m(U') \subset D_{2\widetilde{c}\tau_m^2} \subset D_{\tau_m}, \tag{7.47}$$

by the induction hypothesis, (7.43) and the inequality $2\widetilde{c}\tau_m^2 < 2\widetilde{c}\tau_{m-1}^2 < \tau_m$ (the decreasing of the sequence τ_j and (7.41)). One has $T_{m-1} + \tau_m = T_m < \tau$. This together with (7.46) and (7.47) proves the induction step and (7.45). $\square$

Statement (7.45) immediately implies (7.35). $\square$

Statements (7.26) and (7.27) follow from (7.35), as was shown above. The proof of Theorem 1.26 is complete. $\square$

Proof of Theorem 1.21 with the boundedness of derivatives (Remark 1.22). To show that a Lie group with a Lie algebra satisfying the (strong) Solovay-Kitaev inequality is $\varepsilon'(x)$- approximable (see (1.10)) with bounded derivatives, we replace the previously defined Ω'' by

$$\Omega'' = \{[x, y] \mid x, y \in \Omega\}.$$

Then an analogue of Lemma 7.3 holds true for this Ω''. Afterwards the proof repeats the previous one (of Theorem 1.26) with obvious changes. $\square$

8 Acknowledgments

I am grateful to É.Ghys who had attracted my attention to the problem. I wish to thank him and also É.Breuillard, J.-F.Quint, B.Sevennec, G.Tomanov, E.B.Vinberg for helpful conversations. I am grateful to the referees for very valuable remarks, which inspired major improvements. A significant part of the proof of the Main Technical Lemma was obtained

during my stay at RIMS (Kyoto). Some parts of the paper were written while I was visiting the University of Toronto. I wish to thank both institutions for hospitality and support.

The research was supported by part by Russian Foundation for Basic Research (RFFI) grants 02-01-00482, 02-01-22002, 05-01-02801-NTsNIL_a, 07-01-00017-a.

References

[1] Abels, H., Margulis, G.A., Soifer, G.A., Semigroups containing proximal linear maps. - Israel J. Math., 91 (1995), 1-30.

[2] Arnold, V.I. Dopolnitelnye glavy teorii obyknovennykh differentsialnykh uravnenii. (Russian) [Supplementary chapters to the theory of ordinary differential equations] - "Nauka", Moscow, 1978.

[3] Apanasov, B.N., Gusevskii N.A., Krushkal, S.L., Kleinian groups and uniformization in examples and problems (in Russian). - Novosibirsk: Nauka, 1981.

[4] Breuillard, E., Gelander, T. On dense free subgroups of Lie groups. - J.Algebra, 261 (2003), No. 2, 448-467.

[5] Brown, G. On commutators in a simple Lie algebra. Proc. Amer. Math. Soc. 14 1963 763–767.

[6] Dawson, C.M.; Nielsen, M.A. The Solovay-Kitaev algorithm. Quantum Inf. Comput. 6 (2006), no. 1, 81–95.

[7] Gamburd, A; Jakobson, D; Sarnak, P. Spectra of elements in the group ring of $SU(2)$. - J. Eur. Math. Soc. (JEMS) 1 (1999), no. 1, 51–85.

[8] Gelander, T. On deformations of free subgroups in compact Lie groups. - To appear in Israel J. Math.

[9] Ghys, É., Carrière, Y. Relations d'équivalence moyennables sur les groupes de Lie. - C. R. Acad. Sci. Paris Sér. I Math, vol. 300 (1985), No.19, 677-680.

[10] Goto, M. A theorem on compact semi-simple groups. J. Math. Soc. Japan 1, (1949). 270–272.

[11] Epstein, D. B. A. Almost all subgroups of a Lie group are free. - J. Algebra 19 1971 261–262.

[12] Hirschbühl, R. Commutators in classical Lie algebras. Linear Algebra Appl. 142 (1990), 91–111.

[13] Ilyashenko, Yu.S.; Pyartli, A.S. The monodromy group at infinity of a generic polynomial vector field on the complex projective plane. - Russian J. Math. Phys. 2 (1994), no. 3, 275–315.

[14] Kaloshin, V., Rodnianski, I. Diophantine properties of elements of $SO(3)$. - Geom. Funct. Anal. 11 (2001), no. 5, 953–970.

[15] Kitaev, A. Yu. Quantum computations: algorithms and error correction. (Russian) Uspekhi Mat. Nauk 52 (1997), no. 6(318), 53–112; translation in Russian Math. Surveys 52 (1997), no. 6, 1191–1249.

[16] Lyubich, M. Yu. Dynamics of rational transformations: topological picture. (Russian) - Uspekhi Mat. Nauk 41 (1986), no. 4(250), 35–95, 239.

[17] Nielsen, M.A.; Chuang, I.L. Quantum computation and quantum information. Cambridge University Press, Cambridge, 2000. xxvi+676 pp.

[18] Slodkowski, Z. Holomorphic motions and polynomial hulls. - Proc. Amer. Math. Soc. 111 (1991), no. 2, 347–355.

[19] Sullivan, D.P.; Thurston, W.P. Extending holomorphic motions. - Acta Math. 157 (1986), no. 3-4, 243–257.

[20] Tits, J. Free subgroups in linear groups. - J.Algebra, vol. 20 (1972), 250-270.

[21] Vinberg, E. B.; Onishchik, A. L. Seminar po gruppam Li i algebraicheskim gruppam. (Russian) [A seminar on Lie groups and algebraic groups] - Second edition. URSS, Moscow, 1995.

Poncelet Laboratory (UMI of CNRS and Independent University of Moscow), 11, Bolshoi Vlasievskii Pereulok, 119002, Moscow, Russia

CNRS, Unité de Mathématiques Pures et Appliquées, M.R.,
École Normale Supérieure de Lyon
46 allée d'Italie, 69364 Lyon Cedex 07 France
Email: aglutsyu@umpa.ens-lyon.fr